Aalto University
School of Science
Master's Programme in Information Networks

Lotta Ahonen

Privacy experience – chasing the creepy line

Master's Thesis

Helsinki 30.1.2017

Supervisor: Marko Nieminen

Instructor: Katarina Segerståhl

Aalto University
School of Science

Aalto University School of Science Master's Programme in Information Networks	ABSTRACT OF THE MASTER'S THESIS

Author: Lotta Ahonen

Title: Privacy experience – chasing the creepy line

Number of pages: 54	Date: 30.1.2017	Language: English

Code: SCI3047

Supervisor: Marko Nieminen

Instructor(s): Katarina Segerståhl

This thesis is set to investigate questions regarding individuals experiencing privacy. From the industry point-of-view the question is what type of data and how much of it are people willing to give to organizations. What will make people give their information? In order to answer this question, the phenomenon of privacy experience needs to be studied.

These questions were addressed with a set of interviews. These interviews consisted of trying out a prototype of a shopping assistant application, interview and a short questionnaire. Three different versions of the prototype were used in order to test how freedom of choice and information transparency affect the participants' answers.

Based on the research and literature, what and how much information people are willing to give, are dependent on many different factors. These factors are related to the factors that define privacy experience. They can be divided into contextual and individual factors. These are previous experience, trust, privacy calculus, control, information transparency, type of information shared and awareness. Thus, it is not possible to construct a general guideline to indicate a certain point where people generally stop giving personal information.

Keywords: privacy experience, privacy experience model, data collection, privacy segments, privacy paradox

Aalto-yliopisto
Perustieteiden korkeakoulu

Aalto-yliopisto Perustieteiden korkeakoulu Informaatioverkostojen koulutusohjelma	DIPLOMITYÖN TIIVISTELMÄ

Tekijä: Lotta Ahonen

Työn nimi: Yksityisyyden kokemus – milloin siististä tulee pelottavaa?

Sivumäärä: 54	Päiväys: 30.1.2017	Julkaisukieli: Englanti

Pääainekoodi: SCI3047

Työn valvoja: Marko Nieminen

Työn ohjaaja: Katarina Segerståhl

Tämän diplomityön tavoitteena on tutkia yksityisyyden kokemusta. Teknologiateollisuuden mielenkiinto keskittyy siihen, millaista tietoa ihmiset ovat valmiita antamaan itsestään yrityksien käyttöön sekä kuinka paljon. Mikä saa ihmiset luovuttamaan henkilökohtaisia tietojaan? Jotta näihin kysymyksiin voidaan saada vastauksia, tulee yksityisyyden kokemusta käsitteenä tarkastella perusteellisesti.

Edellämainittuihin kysymyksiin lähdettiin hakemaan vastauksia haastattelujen avulla. Nämä haastattelut sisälsivät ostosavustaja-sovelluksen prototyypin testausta, itse haastattelun sekä lyhyen kyselyn. Prototyypistä testattiin kolmea erilaista versiota. Näiden eri versioiden tarkoituksena oli testata, miten valinnanvapaus ja informaation läpinäkyvyys vaikuttavat testihenkilöiden vastauksiin.

Tutkimuksen ja kirjallisuuden perusteella voidaan todeta, että se kuinka paljon ja mitä tietoja ihmiset antavat, riippuu monesta eri tekijästä. Nämä tekijät ovat samoja tekijöitä, jotka vaikuttavat siihen, miten ihminen kokee yksityisyyden. Ne voidaan jakaa kontekstuaalisiin ja yksilöllisiin tekijöihin. Näitä ovat aikaisemmat kokemukset, luottamus, tietojen luovuttamisen hyötysuhde, kontrolli, informaatin läpinäkyvyys, jaetun tiedon tyyppi ja tietoisuus. Täten voidaan todeta, että ei ole mahdollista konstruoida yleispätevää ohjetta sille, kuinka paljon ja minkä tyyppistä informaatiota ihmiset ovat valmiita luovuttamaan.

Asiasanat: yksityisyyden kokemus, yksityisyyden kokemuksen malli, datan keräys, yksityisyyden kokemuksen eri segmentit, yksityisyysparadoksi

Acknowledgements

Katarina Segerståhl,
Marko Nieminen,
Research participants,
Colleagues,
Friends,
Family.

Thank you all for pushing me forward and nagging to me to get this thing done.

Helsinki, January 30th, 2016

Lotta Ahonen

Table of contents

1. Introduction

"If this is the age of information, then privacy is the issue of our times" stated Acquisti et al. (2015) in Harvard Business Review. Privacy has been getting more and more coverage in the media as well as in the academic world. Activities that once were private, like communication, shopping and dating, leave now digital traces of one's action. (Acquisti et al. 2015) People leave those traces of themselves in the world now more than ever. Over 90 per cent of the data in the world has been created in the past five years (SINTEF, 2013). Sharing data openly can benefit both individuals and the society. Individuals can get more personalized services that serve their needs better. The society can benefit and learn from large interconnected databases. This knowledge can help for example in new medical breakthroughs. But it also has a downside to it. Abuse of personal information is something that is being considered as a constantly growing threat. (Acquisti et al. 2015)

Data collection is nothing new, it has been done for years. Personal data is collected everywhere, in the digital and physical space, whether one realizes it or not. Data is being collected from everything we do, from browser cookies to GPS-location. The tools and methods for analyzing this data have developed vastly in recent years. A good example is the story of Target and their development of personalization. Target had been collecting data about their customers' purchasing habits. Based on this data, they could see certain patterns emerging. One of these findings were that Target was able to tell if someone was pregnant, based on what they were buying and changes in their buying habits. Target sent discount coupons for baby supplies to customers that matched the shopping habits of a pregnant person. One of these happened to be a 16-year-old girl, whose father got rather upset after seeing what type of advertising Target was sending to his teen-aged daughter. After a heated call to Target, blaming them for false advertising, the father had a talk with his daughter. It happened to be so, that the girl actually was pregnant, without her father's knowledge. (Hill 2012)

This story of Target shows how much power companies have over consumers with the help of analyzing traces of data consumers have left behind. Personalization is becoming increasingly important to companies in order to attract more users with personal services and create more revenue. In order to create personalized services, some detailed personal information needs to be collected. On the other side, information collection raises privacy concerns in consumers. If these concerns are not addressed properly, it might lead to refusal of sharing personal information, bad word of mouth or even loss of new customers. Thus, privacy and how consumers experience it, has become a central point of focus in the strategy for companies who operate in the digital era. (Culnan & Armstrong 1999; Awad & Krishnan 2006)

From the consumer perspective, the story of Target can seem a bit scary. How can one company know such things about me without my knowledge? For an individual, keeping track of what personal data is collected and by whom, is getting more and more difficult. Who am I granting permission to use my data and how they actually use it are, questions that consumers face regularly. According to Norberg et al. 2007, this can result in deterioration of the sense of personal privacy. Furthermore, this deterioration leads to decreased willingness to use certain services.

Companies face the dilemma of wanting to provide new personalized and interesting services but at the same time address consumers' concerns about data collection. The question is, when does information collection become uncomfortable to the consumer. Where is the line between creepy and cool?

1.1 Motivation

The motivation for this thesis came from the industry perspective and their needs. Just like the story about Target, also Finnish companies have started to collect very detailed information about their customers. For example, one of the biggest grocery stores has changed their policy so that they can track individuals' purchases on a very detailed level if their loyalty card has been used. This has raised concerns among people. Not knowing what the company will do with all that data is making people uncomfortable. (Yle Uutiset, 2016)

Every company that collects data about their customers is in a way in the privacy business. This has led up to the question about privacy. Who owns the digital information? Who decides how to use the collected information? Data collection enables the development of new personalized services, but it also creates uncertainty about the correct use of the data. Uncertainty, in most cases, feeds the build-up of concerns in people´s minds. Due to this, privacy is the hot question of our times. (Acquisti et al. 2015)

These are question that many modern companies should be thinking about. Companies that are dealing increasingly with personal information and personalized services, see privacy concerns and its effects, such as lack of willingness to disclose personal information, as a major threat to their businesses and companies. (Gurung et al. 2014) This is why the industry of electronic commerce is highly interested in privacy research and solutions to minimize privacy concerns. Another good reason for companies to think about consumer data privacy comes from the EU. In 2018, a new law, General Data Protection Regulation (GDPR), will come into effect that forces companies to make collected personal data available for the consumer (Blackmer 2016).

1.2 Scope

This thesis was done in cooperation with a large Finnish ICT-company Tieto. This affects the scope of the research so that especially business-to-consumer electronic commerce is at the focus of interest as that was the wish from the company. This gives the thesis a close connection to the industry and also provides a bridge between the academic world and business. This thesis should be of interest to both academics and the industry practitioners.

Privacy is a remarkably wide term. Privacy in general has been researched in multiple different fields and meanings. (See Awad & Krishnan (2006) for an extensive list of different types of privacy research.) In order to keep this thesis compact, the scope of this thesis has been limed to address only digital privacy. It has to be mentioned though that the border between physical and digital privacy will be increasingly obscured thanks to the increasing number of different sensors and probes that collect data from the physical human being.

1.3 Objective and research questions

The objective of this thesis stems from the industry viewpoint. As data is being collected from us constantly, either with permission or unknowingly, the question about managing one's own privacy has become a hot topic of the day (Acquisti et al. 2015). What if the situation could be turned around and the power of sharing personal information could be given to the users? If retailers could enhance users' experience about privacy and sharing their data, ease the worry about privacy issues, would people come to their store more often? This chain of thoughts leads up to the main objective of this thesis: how much personal information are people willing to share about themselves, especially in retail context. Thus the main research question (MRQ) is as following:

MRQ: What type of data are people willing to share about themselves in a retail context?

In the process of investigating the area of the main research question, it turned out that how people experienced privacy affected a lot on how they would behave when interacting with a service (Oulasvirta et al. 2014). In order to get to the main objective of this thesis, also the subject of privacy experience needs to be addressed. In the beginning of the research, it came clear that there is not that much holistic research done on the subject. A clear indication of this is the lack of a proper definition of privacy experience. No definitions of privacy experience were found in the process of making this thesis. The subject is mentioned in studies in multiple different fields, but it never seems to be the center of the research. Thus, the supporting research questions (SRQ) are as following:

SRQ1: What is privacy experience?

SRQ2: How can privacy experience be measured?

The main research question will be addressed mostly with an empirical approach but it will also have some backup from literature. As a result, a better understanding will be formed about what information are people willing to share about themselves and why. The supporting research questions will be answered with a theoretical approach by conducting an extensive literature review on the subject. In the end, a construct of the affecting factors of privacy experience will be formulated.

The main research question is dependent on the results of the two supporting research questions. Thus, it will be answered last in this thesis. The supporting research questions will be addressed first by presenting a literature research on the topic.

2. Literature research

The proses of literature research began by looking for articles with the search word "privacy experience". That search word returned only a few results. Thus, the terminology had to be widened. In addition to privacy experience related search words, terms such as perceived privacy, information transparency and information sharing willingness were used. The articles used in this literature research were retrieved from data bases such as EBCO host, Sage Journals, Oxford Journals, Wiley Online Library, Harvard Business Review, IEEE, ACM, Elsevier, MIS quarterly. Google Scholar was used as the main channel of search.

Privacy research is very close to security research. However, security literature is disregarded in this thesis as the main focus is to try to get a better understanding about privacy itself. Articles regarding privacy are taken from multiple different fields as the subject can be viewed from many different angles. social science, marketing, psychology, information technology & management and law.

2.1 Privacy as a field of study

Privacy as a field of study has its roots in the field of law (eg. Schoeman 1984). Particularly consumer privacy, due to its complex nature, has attracted the interest of researchers from several different fields. These fields include social sciences, information systems, marketing, psychology, information technology and management, and law. This study focuses mostly on literature from the field of information systems but also includes literature from the fields of marketing and social psychology. (Brandimarte et al. 2012)

In the field of information systems, privacy has been studied from different structural levels. In their studies, Skinner et al. (2006) have identified three different levels of information privacy: individual, group and organization (Skinner et al. 2006). Smith et al (2011) added one level on top the previous three in his research: societal (Smith et al. 2011). Research usually focuses on the individual and organizational levels. This study mostly focuses on the individual level of privacy as that is where the fundamental experience happens and that is also the level the majority of information privacy research is conducted (Bélanger & Crossler 2011). The societal level is also taken into consideration especially when discussing how previous experiences affect privacy experience.

Because of its complexity, privacy as a concept can be divided into various dimensions. Leino-Kilpi et al. have compiled a four-dimensional model of privacy that is based on findings in previous research (Figure 2-1). This model presents the social, physical, psychological and the informational aspects of privacy. The social dimension is about being able to control interactions with other people as well as the effort to control social

contacts. The physical dimension refers to the right of personal space and territory and has its background in research of live animals. The psychological dimension concerns the ability to control what thoughts or intimate information one shares and with whom. It has been said that psychological privacy has to do with personal growth and self-identity. Informational privacy is the dimension of privacy that has emerged most recently as computers and their ability to store and process data has become a large part of everyday life. Informational privacy concerns the ownership and distribution of one's personal information. (Leino-Kilpi et al. 2001) This thesis focuses on the informational dimension of privacy.

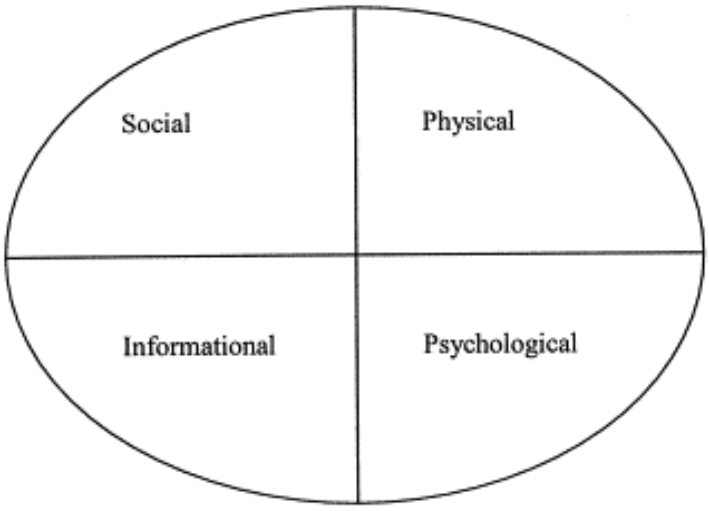

Figure 2-1:Dimentions of privacy (Leino-Kilpi et al. 2001)

2.2 Definitions of privacy

As mentioned earlier, privacy has been studied from multiple different viewpoints. This has led to the fact that the term privacy does not have one solid definition. Different disciplines have created their own definitions of the concept of privacy (Whitley 2009). Next, some definitions from different fields will be presented and discussed.

Within the field of social sciences, privacy is mostly described as control (Xu et al. 2008). It has been said to be "control over personal information flows" (Brandimarte et al. 2012) and "privacy represents the control of transaction between person(s) and other(s), the ultimate aim of which is to enhance autonomy and/or minimize vulnerability" (Margulis 1977).

In literature from the field of law, the definitions of privacy focus more on individual rights but also brings up the importance of control. In law, privacy is also seen as a right or an entitlement (Xu et al. 2008). For example, Solove has listed the following as the core attributes of privacy: the right to be left alone, limited access to the self, secrecy, control over personal information, personhood, and intimacy (Solove 2002). Already a century earlier Warren & Brandeis described privacy as "The right of an

individual to be left alone and able to control the release of his or her personal information" (Warren & Brandeis 1890).

Information technology also includes the acquisition of personal information in the definition. Alan Westin, one of the first to research how people experience privacy, defines privacy as the "ability of the individual to control the terms under which personal information is acquired and used". (Westin 1968) Another more recent definition is "individual's ability to control the terms by which their personal information is acquired and used" (Chellappa & Sin 2005). The most recent definition of privacy comes from the information technology field from Betsy Masiello in the year 2009. She defines modern time privacy as "the right to not be mischaracterized, unsettled, or surprised by what personal information and communications about you are publicly available on the Web". (Masiello 2009)

The example from the field of psychology does not differ dramatically from the definitions of the previously presented fields. Stone et al. describe privacy as "the ability of the individual to personally control information about one's self". (Stone et al. 1983) In general, privacy is seen as limited access or a state of isolation in the field of psychology (Xu et al. 2008).

If we look at the development of the definition of privacy, regardless of the area of study, we can clearly see how the focus has shifted from protecting the physical self to protection of personal data and information and the digital self. This presumably is connected with the development of modern digitalized society. We now see that we also have a digital me besides the physical me. With the rapid development of sensors, soon we cannot separate completely the physical and digital me as data is also being collected from our physical doings. This thesis has the emphasis on the later definitions of privacy, focusing on digital privacy.

Drawing a conclusion from the above presented different definitions of privacy, it is clear that privacy is all about individual's control over their personal data, be it collected digitally or physically, and the ability to see and control the use of it. However, this also supports previous observations from Smith et al. (2011) that privacy really does not have one single definition that would cover all different areas of research.

2.3 Privacy segments

In the mid 90's a survey was developed by Louis Harris & Associates and Alan Westin to measure and segment the public based on attitudes towards privacy issues. The goal was to map how people felt about privacy. Three segments were identified: privacy fundamentalists, privacy unconcerned and privacy pragmatics. *Privacy fundamentalists* were described as people who are very skeptical about privacy issues and only trust on legal and regulatory measures when dealing with privacy. The *privacy unconcerned* were

described quite the opposite from privacy fundamentalists. They are ready to provide any type of personal data to governments and businesses. The biggest group identified was *privacy pragmatists* who will decide case by case if they trusted the organization with their personal data or not. (Westin 2003) In different studies, the privacy fundamentalists usually cover 25 per cent of people, pragmatics around 56 per cent and unconcerned the final 19 per cent of people (Jai & King 2015; Kumaraguru & Cranor 2005; Westin 2003).

The 2002 Harris report provides the following representative descriptions of the different privacy segments:

"**Privacy Fundamentalists**: At the maximum extreme of privacy concern, Privacy Fundamentalists are the most protective of their privacy. These consumers feel companies should not be able to acquire personal information for their organizational needs and think that individuals should be proactive in refusing to provide information. Privacy Fundamentalists also support stronger laws to safeguard an individual's privacy." (Krane et al. 2002)

"Privacy Pragmatists: Privacy Pragmatists weigh the potential pros and cons of sharing information, and evaluate the protections that are in place and their trust in the company or organization. After this, they decide whether it makes sense for them to share their personal information." (Krane et al. 2002)

"Privacy Unconcerned: These consumers are the least protective of their privacy – they feel that the benefits they may receive from companies after providing information far outweigh the potential abuses of this information. Further, they do not favor expanded regulation to protect privacy." (Krane et al. 2002)

Westin developed a Privacy Segmentation Index based on over thirty surveys between 1978 and 2004 (Kumaraguru & Cranor 2005). Even though the Privacy Segmentation index has been very influential, a lot of doubt towards the index has been presented. One of the reasons is that researchers have not been able to establish a salient correlation between the privacy segments and people's actual, contextual behavior (Consolvo et al. 2005; King & Hoofnagle 2008; Malheiros et al. 2013).

Kang et al. 2015 discovered in their research that people's privacy concern levels and respectively actions change based on the situation. In a familiar situation with a previously used service, a user might be very privacy unconcerned, whereas in a new situation with a new service the same user might be a privacy fundamentalist. Hence, Westin's privacy segments cannot be concerned to be comprehensive in every situation as they measure general privacy-related attitudes. They lack evidence between presumed behavior and actual contextual behavior (Fishbein & Ajzen 1977). Privacy

preferences and concerns should be studied in a specific context or circumstance as they are found to be hard to generalize. (Iachello & Hong 2007)

Westin's privacy segments have also been criticized for not taking into consideration the complex nature of people's privacy experience, especially in the modern digitalized world. Privacy needs are becoming increasingly complex due to the intricate nature of the modern information society (Woodruff et al. 2014). Researchers have shown that privacy is an intricate matter where concerns do not always align with behavior. Westin's privacy segments do not actually address this aspect at all. The distinct gap between people's privacy attitudes and actions, referred to as privacy paradox (see 2.4), is not addressed in Westin's privacy segmentations. (Woodruff et al. 2014) Instead of extremely general studies, the researchers are calling for more detailed research about privacy attitudes and behavior in order to get to the bottom of privacy experience. (Kang et al. 2015; Woodruff et al. 2014)

Westin's privacy segments have been studied further later on. For example Smith et al. (1996) have done additional research, where they identified four subscales, that supplement the already identified three privacy segments. These are called concern for information privacy (CFIP). They are concerns about collection of personal information, processing errors, further use of personal data and improper access to the information. The CFIP consists of 15 items and is nowadays widely known as an instrument used in factor analysis for information privacy constructs. (Bélanger & Crossler 2011) A few years after CFIP was developed, internet user's information privacy concerns (IUIPC) was developed, based on the CFIP. The IUIPC contains only three dimensions: control, awareness and collection. It is also less used than the original CFIP measures. (Bélanger & Crossler 2011) These measurements will be discussed in more detail in section 2.6.

2.4 Privacy paradox

When it comes to online privacy, people tend to be quite worried about privacy on a general, societal level. However, this did not always translate to action on the personal level. Even if people were aware of privacy threats and worry about them, they might have completely neglected taking any actions in order to enhance their privacy. People have the tendency of thinking that "it won't happen to me". (Cho et al. 2010; Gross & Acquisti 2005) This phenomenon is known as privacy paradox. More precisely, privacy paradox is described as the distinct difference between individuals' privacy-related attitudes and their actual behaviors regarding privacy protection and information disclosure (Norberg et al. 2007).

The phenomenon of privacy paradox has been recognized and studied multiple times and many have reported to have found differences between individuals' attitudes and actual behavior. (Acquisti & Grossklags 2005; Norberg et al. 2007) It is widely acknowledged in theoretical research across different fields that usually person's

intentions lead to behavior that goes together with the intention. However, this does not hold true in the context of privacy related behavior. (Bélanger & Crossler 2011) This fault has been proven in empirical research. For example, Kang et al. (2015) have studied privacy in the context of loyalty card usage and found that their participants were unsure about how their data would be used by the store. However, the participants did not take any actions in order to ease the feeling of uncertainty, and continued using the loyalty cards.

Why do people's intentions and actions not match and why exactly this privacy paradox phenomenon exists is still something research has not been able to completely answer (Norberg et al. 2007). One advice for researchers has been given: not to assume that intentions lead to behavior when conducting research on privacy (Bélanger & Crossler 2011).

2.5 Defining privacy experience

Going through literature related to privacy experience, one thing became clear. There is a gap in the research field of privacy experience. Especially in the field of Information Systems research, different studies focus on different aspects of the concept of privacy experience. There are no holistic studies made about the privacy experience as a whole. (Bélanger & Crossler 2011)

There is no proper definition available for the term privacy experience. One article was found, by Betsy Masiello (2009) that solely focused on discussing what privacy experience actually constructs of. Due to the very limited amount of material on the definition of privacy experience, a deeper look into the subject is definitely needed. This section will focus on identifying factors that affect privacy experience and thus contributes to the research and definition of privacy experience.

Since literature on privacy experience is quite limited and diffused, the search terminology needed to be extended beyond the term "privacy experience". Thus, literature for the purpose of defining privacy experience consists also of material that is sought with search terms such as "perceived privacy", "privacy research" and "experiencing privacy".

The research on individual's experience of privacy seemed to be divided into two fields. Several previous studies focused either on individual or contextual factors of privacy. Rarely have these two views been combined in order to create a more holistic view about privacy experience, even though it has been stated that neither the contextual nor the individual factors alone explain the experience an individual has about privacy (Martin & Shilton 2015).

The interest in this thesis is in understanding privacy experience in a comprehensive manner. Thus, a model of privacy experience is compiled that combines both the

individual and the contextual factors. The purpose of this model is to get a better overview of the different blocks that privacy experience builds on and to test them in practice with the help of qualitative research and interviews. However, this thesis is not able to address all of the factors in the qualitative research phase, hence we are focusing only on a couple relevant ones. The most relevant factors are chosen based on how much emphasis they got in previous studies and on the relevance to the concept of personalized shopping applications as that is the most interesting viewpoint of the industry related to this thesis.

The different factors that have been selected to put in this model are selected based on previous literature that was acquired with search terms such as "privacy experience", "perceived privacy", "privacy research" and "experiencing privacy". The articles found with these queries were analyzed to see if they 1) tried to address the complex nature of privacy experience in some way 2) had a way to measure it. The emphasis was on the first requirement. In the end, altogether 40 articles were selected for further analysis in order to construct a holistic outlook on privacy experience. From these articles, some already existing models of how privacy affects behavior were found. For example, Liu et al. (2005) tested a privacy-trust-behavioral intention model. In their model privacy consisted of four dimensions: notice, access, choice and security. Smith et al. (1996) identified four dimensions that affect the individuals' concern for privacy: collection, errors, secondary use and unauthorized access. By combining these different factors identified in previous research, the model of privacy experience formed.

The model created in this thesis is heavily influenced by Martin & Shilton's (2015) previous research. In literature research for this thesis they were the first ones that were found that commented on the division to individual and contextual factors of research in privacy. They use a model where individual and contextual factors are combined and seen as equal influencers to privacy experience. This combined model is used as a frame for creating the model in this thesis. All of the emerged factors effecting privacy experience have been divided into either individual or contextual factors (Table 2-1). The division has been made based on if the factor is something the service provider can directly effect through the service or is it something that the user brings to the situation. Some of the factors could be placed on both sides but in the name of simplicity, the more obvious side has been chosen.

The next seven subchapters will go through more thoroughly the different factors and their effects on privacy experience.

Individual factors	Contextual factors
Previous experiences	Awareness
Trust	Type of information shared
	Information transparency
	Control

Table 2-1:Factors effecting privacy experience

2.5.1 Contextual factors

The contextual factors awareness, type of information shared, information transparency and control, are factors that users' have no saying to. They are controlled by the decisions of service providers. The contextual side of privacy experience sees it as a phenomenon that is contextually dependent, based on the content and context of the service (Martin & Shilton 2015).

Awareness

Awareness in the context of this thesis refers to an individual's understanding of what personal information is collected or shared and with whom, who is the receiver of the information and how the information is used. Awareness can also be technology awareness, awareness about privacy protection and fair procedures on websites.

Low level of user awareness is often linked with an increased amount of privacy concerns the user experiences on the internet. Not knowing who automatically collects what type of personal information or how it is being used, creates uncertainty in most people. This can present itself in various ways, such as insecure feeling about disclosing personal information, uncertainty where to share information safely or the fear of uninformed third-party use of personal information. (Acquisti et al. 2015; Culnan & Armstrong 1999) Low level of user awareness is also said to be a reason for people not following common security advices or ignoring basic privacy features online. If the risks of privacy violations and their consequences are not understood by the users of online services, how could one expect them to be able to understand the benefits of privacy protection tools and procedures? (Das et al. 2014)

As the level of technology awareness grows, the individuals' amount of privacy concerns usually lowers and confidence grows. This is due to the fact that people become smarter in their online actions. They begin to understand what are the appropriate online habits and how information from different places connects together. (Gurung et al. 2014) Moreover, risen levels of awareness lead to understanding ways to use more protective behavior and technologies against privacy threats (Dinev & Hu 2007).

Just by being aware of data collection and the surveillance happening online, one might gain a more understanding attitude towards the reasons behind data collection. Allen et al. (2007) gives an example of this from the physical world, in a case of workplace monitoring. When the surveillance was done in secret from the employees and it was discovered, little understanding was given towards the act of surveillance. But if the employees were informed beforehand about the surveillance and the reasons behind it, they had a more positive attitude towards the surveillance. (Allen et al. 2007)

On the other hand, increased awareness might have the complete opposite effect. As the level of awareness grows, so does the understanding of what is not understood. Being aware about information collection but being unsure or unable to do anything in order to prevent the information collection creates an uncertain feeling about the online world (Kang et al. 2015). Understanding more about the complexity of the use of personal information and being more aware about the fact that in some services it is not that clear where the information ends up, raises privacy concerns. (Gurung et al. 2014; Liao et al. 2011)

All in all, it looks like it is good to have some level of awareness in order to be able to operate safely in the online world without being entirely petrified about every single action taken. However, when the awareness reaches a certain level, privacy concerns start to build up again due to gained understanding about the unknown and the possible risks lurking around the internet. If the media coverage of security and privacy breaches starts to be regular, awareness and knowledge of the problems are likely to increase among people. Awareness has a clear effect on how people experience using the internet and making transactions there. This is why awareness is also an important factor in the definition of privacy experience.

Type of information shared

Individuals value some types of data differently than others, especially when it comes to sharing it. Different types of data have different sensitivity levels that affect the willingness to share it to other stakeholders. (Malheiros et al. 2013) Thus the type of data will also have an effect on the privacy experience. This chapter will discuss the different sensitivity levels between different types of data and their possible implications to data sharing and eventually privacy experience. It is based on research from different fields, including marketing, public policy, privacy research in computer science and business. The detailed division of different data types and their sensitivity levels can be found in Table 2-2.

The least sensitive type of data seems to be data that is frequently asked from individuals, such as name, age, gender and email (Malheiros et al. 2013). It could be that people are getting used to providing this type of information when browsing the internet and thus they do feel like this type of data is not that sensitive to give out. Research in marketing has found out that education and other demographic

information are also something that individuals are fairly willing to give out (Phelps et al. 2000). (Horne & Horne)

Table 2-2: Sensitivity levels of different types of information

	Most sensitive	Medium sensitive	Least sensitive
Marketing (Horne & Horne 1998)	medical, financial, and family information		product and brand consumption, media usage behavior
Public policy (Phelps et al. 2013)	financial information and personal identifiers, annual household income, the kinds of credit cards they possess, their social security numbers, or their two most recent		demographic and lifestyle information, two favorite hobbies, age, marital status, occupation, or type of
Privacy (Malheiros et al. 2013)	illness, annual income	medical and financial data	gender and education
Business (Rose et al. 2012)	social network posts, health records, financial data, credit card data	past purchases, purchase plans, media usage, location	age group and gender, opinion on products, name and email, interests
In general	Financial and health records. Data that is not easily available to the public, or that will reveal personal and sensitive facts about the		Gender and age, all very simple information

Individuals are most hesitant to share their medical and financial data, or any other type of data that could be identified and linked to them. This type of very sensitive data is usually consisting of credit card information, social security numbers, other financial information and medical records (Phelps et al. 2000; Malheiros et al. 2013).

One thing that seems to have changed over the years is how people view sharing their past purchase history. In research done in the 1980's, product and brand consumption and media usage behavior were seen as not that sensitive information to share (Horne & Horne) whereas nowadays people seem to value this information as more sensitive (BCG). It could be that the more visible use of this data by companies, e.g. targeted marketing, have an effect on this matter.

All in all, people are willing to give away some personal information in order to be part of the modern society. People would be disappointed if they would not have the option to utilize services provided by the availability of their personal information. Finding the right balance between what to share with whom is the key. (Phelps et al. 2000)

(Information) transparency

Transparency in this context is described as giving access to the person being surveilled about what is the identity of the quarter collecting information, what type of information is collected, how it is going to be used and in what practice (Oulasvirta et al. 2014; Awad & Krishnan 2006). It has been noted that more transparency about the intentions of the usage of the collected information will have a decreasing effect on the level of privacy concerns as it reinforces trust towards the information collector. If the information collector is seen to have nothing to gain by misusing the collected information, trust towards that collector is increased, thus creating a more pleasant experience. (Oulasvirta et al. 2014)

In previous, very narrow, privacy experience literature, transparency and choice has been seen as the two founding pillars of privacy experience. This is due to the fact that together transparency and choice establish a suitable environment for collecting data with informed consent (Masiello 2009), which again creates a pleasurable privacy experience to the user. Later research has confirmed the importance of transparency in creating a usable privacy experience (Oulasvirta et al. 2014; Morey et al. 2015; Cavoukian 2009).

Transparency also leads to an increase in the individuals' willingness to share personal information as the individual is able to see and assess the collected information and the possible use of it. (Oulasvirta et al. 2014; Morey et al. 2015) In previous research it has turned out that people would be willing to give personal information about themselves if the use of the information would be clearly stated. (Culnan & Armstrong 1999) However, there is another side to transparency. Individuals, who value transparency

the most, are usually the ones that do not want to be profiled. It is suggested that information collectors should also put more emphasis on being transparent about the increased value, the benefit that the collected information can provide, and also the possible risks (Awad & Krishnan 2006; Masiello 2009). As an example, Awad & Krishnan (2006) take the difference between personalized advertisement and personalized services. People are more willing to give out personal information for a personalized service than for personalized advertising as it more clear to them that the personalized service will bring more value to them than just personalized advertising. Giving personal information to advertising has been seen riskier as it is not always transparent what happens to the information afterwards (Awad & Krishnan 2006).

In conclusion, it could be said that transparency is the factor that allows the users and companies to be more on the same level, as it reveals intentions, actions and information to all parties involved. It balances the authority relationship and tries to achieve equality. (Zarsky 2004).

Control

Control, in the context of privacy and thus this research, is defined as "limiting what personal data is made available to others". (Whitley 2009) More precisely, this section focuses on the control of the individual over his or her own personal data and the possibility of choice.

Lack of control over personal information has been identified as one of the most important aspects in the privacy discussion regarding online environments and how to create an enjoyable experience. (Whitley 2009; Culnan 1995) One of the crucial parts of how consumers experience privacy online is the consumers' ability to control their actions on the website. If the user feels that he or she is in control of the personal information and can choose what to share, it can reduce the feeling of privacy risks associated with the usage of online services and giving away personal information. This can lead into resolving some privacy-related anxiety (Hoffman et al. 1999; Phelps et al. 2000; Gurung et al. 2014). Assuming that people feel that they have control over the possible future use of their personal information collected, people will consider information collection less invasive and perceive the service more trustworthy. (Culnan & Armstrong 1999; Liu et al. 2005) Some examples of this type of actions are deciding what personal information the person wants to share and how the shared information is going to be used in the future. (Chen & Rea 2004; Liu et al. 2005)

However, people usually take more risks when they feel that they are in control. For example, people feel that it is safer to travel by car than by plain. One reason for this is that by driving a car, people are more in control of what happens on the journey than on the plain, where they need to trust that the pilot keeps them in the air. The same phenomenon can be seen with cars and seat belts. People assume and trust that

seat belts will protect them from major injuries, hence they start to drive more recklessly. (Brandimarte et al. 2012)

The feeling of being in control also interrelates with optimism bias. The more a person feels that he or she has control over a situation, the more his or her optimistic bias grows for that situation. This creates an illusion of being invulnerable and comparatively superior to others. (Cho et al. 2010)

Being more careless when having the feeling of control also applies to collecting and sharing personal information, especially when the control is given explicitly. The more control people feel that they have the more they will start to trust the information collector. At the same time, people are willing to take more risks and share more personal information. (Brandimarte et al. 2012; Acquisti et al. 2015) A balance on the amount of control and what to actually control is also required. When managing personal information becomes a burden, it loses its meaning and might even create new privacy risks (Masiello 2009).

2.5.2 Individual factors

The individual factors previous experience and trust, are factors that are more part of the individual interacting with a service. The individual side of privacy experience sees privacy experience more as a general feeling rather than being tied to a certain situation. (Martin & Shilton 2015)

Previous experiences

It has been widely acknowledged that people constitute mental images about the world based on their previous experiences. People combine previous experiences and assume what will happen in the future based on them. This phenomenon has been studied on two different levels, either on the personal level or on the societal level. Here the personal level refers to the person's own experiences, whereas the societal level refers to others' experiences. (Cho et al. 2010) In this section, both approaches will be discussed at the same time as they are closely related to each other.

As stated in the ISO-standard's definition of user experience, personal previous experiences have an effect on how a service or a system is experienced. This notion can also be adapted to experiencing privacy, as it can be seen as a derivative from user experience. (ISO 9241-210:2010)

People have the ability to distinguish between the two levels of experiences. These are the societal level that include the experiences of other people and the personal level which relates to past personal experiences. People tend to believe more in their own experiences than in others'. Individuals' own experiences are much more powerful than others'. However, if people do not have much prior experience about a certain

situation, they do take others' experiences more into consideration. (Cho et al. 2010; Kang et al. 2015)

People have the tendency to believe that "it will not happen to me" and thus give little attention to others' experiences. This is called **optimistic bias**. Optimistic bias is when individuals believe they are less likely than others to encounter negative things, such as heart attacks or car accidents. This applies also in the context of online privacy. People tend to believe that they are less vulnerable to encounter online privacy breaches than others. Hence, their perception of privacy will change and their privacy related behavior will become more careless increasing the risk of actual privacy violations. Internal beliefs and prior experience have a notable effect on creating a gap between the societal and personal levels and thus widening or narrowing the gap of optimistic bias. (Cho et al. 2010)

The context of use also needs to be taken into consideration. In Awad & Krishnan's (2006) research, they compared two alternative contexts, personalized services and personalized advertising. In the case of personalized services, where people are able to see that they can benefit from sharing personal information, the possible previous privacy invasions do not have a significant effect on the experience. The benefits outweigh the potential risk of privacy invasion. Whereas in the case of personalized advertisement, the risk of violating the use of personal information is seen bigger than the actual benefits gained and thus possible previous privacy invasions have a significant effect on the experience. (Awad & Krishnan 2006)

Trust

Trust is a complex factor in privacy research as it has many meanings and interpretations. It can be a social phenomenon reflecting many different (e.g. behavioral, technological and social) aspects of interactions. Trust can also be part of a person's trait, or a part of his beliefs towards something. (Liu et al. 2005) Trust can also be described as a feeling of security and confidence towards someone or something (Gefen et al. 2003).

This thesis will mostly focus on trust towards organizations and companies when people engage in electronic commerce. In this context, trust has been described as individuals' willingness to be vulnerable and belief that the companies that they are engaged with will not break their trust (Gurung et al. 2014; Martin & Shilton 2015). Different fields of studies have very similar interpretations of trust in the context of electronic commerce. For example, in marketing literature, trust is defined as "a willingness to rely on an exchange partner in whom one has confidence" (Schoenbachler & Gorden 2002). In information systems literature trust is seen as "a belief that one can rely upon a promise made by another" (Pavlou 2003).

It has been noted that trust is an important determinant when it comes to consumer behavior in electronic commerce and the prevention of building up privacy concerns. It has been argued that the lack of trust in online services is the first and foremost reason that prevents companies to grow in the market. This is due to the fact that individuals who do not trust a company do not engage into any kind of relationship with the company let alone provide personal information to that company (Hoffman et al. 1999). Creating trust between an individual and a service has a decreasing effect on privacy concerns and that can lower the individual's barrier to interact with a company. The more people are interacting with a company, the bigger is the probability that they will give personal information to the company, especially if their trust can be increased. (Gurung et al. 2014) Increased trust also makes it possible for the individuals to accept possible uncertainties and risks when it comes to making financial transactions (Chellappa 2002).

The same way as the feeling of control increases the individuals' willingness to share information, the feeling of trust also raises the readiness to publicly share personal information. When a certain level of trust has been established towards an organization or a company, individuals are more likely to be more open towards that organization or company and trust personal information to them in exchange for something. (Schoenbachler & Gorden 2002; Chen & Rea 2004; Dommeyer & Gross 2003; Naresh K. Malhotra et al. 2004) Privacy invasions and breaches cause concerns between the individual and the service and thus decrease trust between them, making the individual more careful about how they interact with the service and what information they provide (Chen & Rea 2004).

When a consumer is interacting with a new online service, especially in electronic commerce, acquiring the initial trust is very important. Building trust online differs from building trust in the physical world. Vendor / customer relationships can generally be described as distant and impersonal (Chen & Rea 2004). In a situation where the buyer and the seller barely know each other, initial trust creation is crucial. In the physical world a store can be judged by its looks, neighborhood, customer service, presence of other customers and size. (Gurung et al. 2014) All of these cues are missing in the digital world. Hence, other means are to be used. Initial trust creation relies on how the customer perceives the reputation of the new service and the quality of their online presence (McKnight et al. 2002). These can be built and enhanced by having a familiar and usable interface that evokes trustworthiness and links positively back to previous experiences with similar services. Showing that the service is secured and that possible transactions take place as expected will increase trust. (Gurung et al. 2014; Culnan & Armstrong 1999) A positive effect on trust is also built through giving the users control and educating them about the purpose of the service, making them aware about the intention of use. (Gefen et al. 2003; Morey et al. 2015)

The most important factor is that the user believes and trusts that the company in question has nothing to gain by cheating or misusing the provided personal data. The user is usually most concerned about what is the data going to be used for and who can access the data, the intention of use being the more important one. Being aware about the possible uses of personal data, privacy concerns can be lowered. This requires transparency from the service providers. (Oulasvirta et al. 2014; Gefen et al. 2003; Ackerman et al. 1999)

Altogether, it can be argued that trust is closely related to all of the abovementioned factors in this section: awareness, control, transparency, previous experiences and type of information shared. Thus, trust has a significant effect also on the privacy experience in general. Trust is the factor that ties together all of the different factors affecting privacy experience.

2.5.3 Summary of the privacy experience model

In the current version of the model (Figure 2-2), there are no indicators of how much each factor affects privacy experience. That is left empty on purpose since it cannot be clearly measured as different factors have a different effect on the experience depending on the context (Martin & Shilton 2015). No relations between different factors have been visualized as it is not the main focus of this study. It is indisputable that the different factors overlap and influence each other.

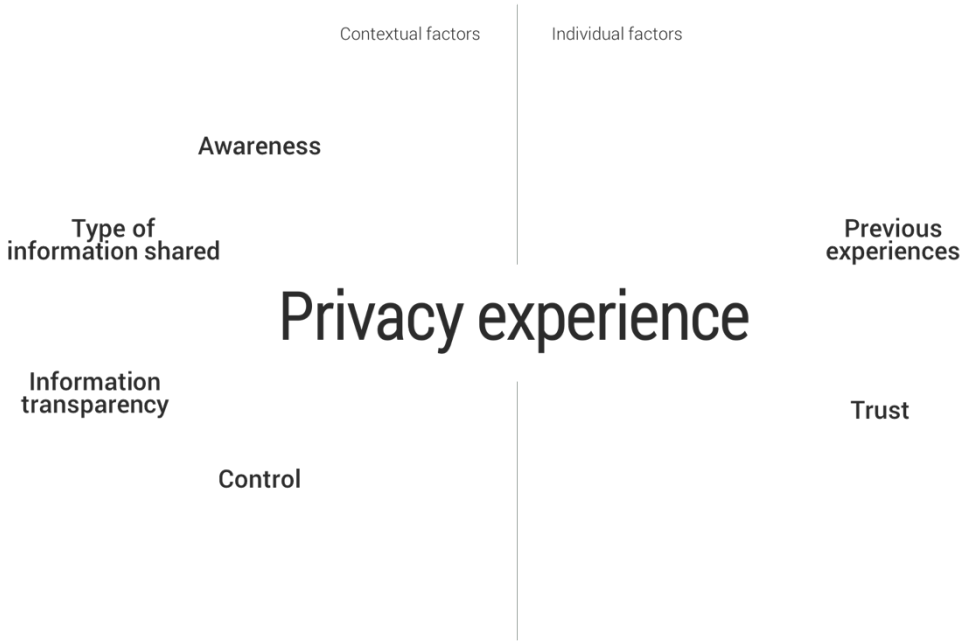

Figure 2-2: The privacy experience model

It has been argued that trust is the main factor that affects privacy experience in the most comprehensive way (Schoenbachler & Gorden 2002). Based on the literature research it seems like all of the factors are connected in some way with trust.
Based on the literature review it cannot be determined if trust has a bigger effect on the privacy experience or the other way around. These causal relationships would need further research.

2.6 Measuring privacy experience

Based on the literature research collected in the previous sections, the most common way to study and measure privacy experience has been executing some type of questionnaire with a large number of participants. Typically, these have been online questionnaires. The collected data is most of the time quantitative and different factor analyses will be performed in order to reveal causal relationships, variances and reliability of the results. Very few of the previous research used any type of qualitative methods, even as one part of their research. This thesis addressed this shortage of qualitative research by using interviews as one of the main methods of empirical research. It was also more suitable for this thesis to focus on qualitative data as the number of participants is quite small as quantitative methods would not have given reliable results.

In addition to Westin's privacy segments, four different methods for measuring and modeling privacy experience were found in the literature research. These were Concern For Information Privacy model (CFIP), Internet Users' Information Privacy Concerns scale (IUIPC), Private Consumer Information Cost (PCIC) and Privacy Concern Scale (PCS).

Both CFIP and IUIPC were derived from Westin's privacy segments (Bélanger & Crossler 2011) and are the most used instruments used in individuals' privacy concern research (Bélanger & Crossler 2011). The Concern For Privacy model (CFIP) was one of the first instruments to be developed to measure individuals' privacy concerns especially in the field of information privacy. It was developed by Smith et al. (1996). This model used 15 items from four different aspects regarding privacy concerns: collection of data, errors in the data, secondary use of data and improper access (Oulasvirta et al. 2014). The CFIP model has mostly been used in the context of offline marketing (Naresh K Malhotra et al. 2004).

The IUIPC scale was developed from CFIP (Bélanger & Crossler 2011). The IUIPC was based on three factors: collection, control and awareness of privacy practices. It was also directed more to the use of online marketing and privacy research in that field. IUIPC took the CFIP to the world of internet (Naresh K Malhotra et al. 2004; Xu et al. 2008)

The Privacy Concern Scale (PCS) was created in 2007 by Buchanan et al. The PCS was a developed iteration of the IUIPC, that would keep privacy measuring up to the speed with the ever changing online world. PCS is most commonly utilized in activities for the online environment, such as registration, e-commerce and emails. (Woodruff et al. 2014)

2.7 Summary

The privacy experience model created based on the literature review, was mostly based on quantitative research. Measuring factors of privacy experience were also based on quantitative methods, mostly different types of questionnaires. However, as the objective of this thesis came from the industry, and this thesis is done in order to support the development of digital services, it is not always possible to get a suitable amount of answers in order to use quantitative research methods. Thus, the research methods used in this thesis were qualitative methods that were based on the quantitative methods.

3. Research methods and process

This section describes the methods used in the empirical part of this thesis. The main method of empirical research, interviews with a questionnaire, was selected based on the literature research, mainly by Culnan & Armstrong (1999). The purpose was to find a previously conducted study in the context of retail that could be replicated in the scope of this research. This would ensure that the empirical research made for this thesis would be academically relevant and meaningful. The empirical research seeks to answer the first research question of What type of data are people willing to share about themselves in a retail context?

In selecting the participants for the interviews, purposeful intensity sampling was used. Purposeful sampling is a qualitative research method that focuses more on the quality of the sample rather than the quantity. "Purposeful samples be judged on the basis of the purpose and rationale of each study and the sampling strategy used to achieve the study's purpose." (Patton 1990)

Purposeful intensity sampling suits this research well as there was prior knowledge from the industry about who would use personalized shopping services. Having the extreme ends of the population in the research would not bring any added value to it as the extreme ends are not in the target group. As the research focuses on a fairly new topic we want to rule out the extreme ends and focus more on the core of the question. At the same time, we don't want to narrow the target pool of participants too much as it might prohibit us from making any reasonable conclusions. That is why it is critical to choose a suitable sample to shed light upon the phenomenon. As Patton puts it, purposeful intensity sampling is used to "…seek excellent or rich examples of the phenomenon of interest, but not unusual cases" (Patton 1990).

Certainly, purposeful intensity sampling also has a few downsides to it. Using this method can induce researcher bias as it is heavily relying on the researcher's own judgement. This can be overcome / reduced to minimum with planning the sample carefully beforehand and having very clear criteria to comply with. Another drawback is that due to the selective nature of the sample it might not be generalizable. (Patton 1990)

3.1 Internet questionnaire

As section 2 demonstrated, peoples' privacy experience depends on various factors. One of the most important ones is trust towards the information collector. The purpose of the first part of the empirical research was to look into who people trust with their personal information and why. This is done in order to get an overview of the current attitudes towards privacy in general, before diving deeper into the area of retail with the second empirical research.

The questionnaire consisted of five questions considering to whom are people willing to trust their personal information to and why, what type of information would they be willing to share and do they trust their own competences in keeping their private information safe. Two of the questions were multiple choice questions, where the participant could choose multiple answers. The other three questions were open questions. The detailed questions can be seen in Appendix A.

The questionnaire was posted in social media (Twitter and Facebook) in December 2015 and it was collecting answers for a month.

3.2 Interviews with a prototype

The empirical research in this thesis consists of people trying out a prototype of a personalized shopping application, followed by an interview and questionnaire after the use of the application. The participants were encouraged to think aloud throughout the whole process in order to catch all of their doubts and wonderments. The purpose of this combination was to first find out what type of information are people willing to give out in the context of retail and electronic commerce, and secondly try to map out people's general attitudes towards data collection and personalization, in order to answer the first research question. All of the questionnaires and interviews were done in Finnish as it was the native language of the participants.

The research was carried out in public spaces, such as cafes and shopping malls in order to create an authentic feeling and atmosphere about the possible use scenario of the personalized shopping application. The tests lasted between 15 and 30 minutes, depending on how much the participant was willing to talk. In the beginning, the participants were given a scenario that they had to adapt to their thinking. The participants had to imagine a situation where they were looking for a new jacket for the spring season. They had to think of what type of jacket they were looking for and for what purpose. After presenting the scenario, the prototype was introduced. It was explained that the application was still in the prototype-phase and that is why it was not looking that nice. The participants were told that they could use a new application that could help them in selecting a new jacket. All they had to do was to answer a few questions about themselves and their current situation and the application would provide them with suggestions about possible fitting jackets and stores where to buy them. It was also mentioned that the information they will provide would be sent to multiple retail stores.

After the prototype, the participants were interviewed and finally given a questionnaire to fill out. The two phases usually mixed together as the questionnaire questions worked also as conversation starters.

3.2.1 Prototype

The prototype developed for this thesis was created with a questionnaire tool called Typeform (Figure 3-1). The original plan was to create a customized application for the prototype, but due to time limitations a questionnaire tool was used. The questions in the prototype were based on requirements from the industry and on a previous research conducted by Spiekermann et al. in 2001. This specific research was chosen based on three criteria. First, the research addressed the topic of experiencing privacy. It also dealt with the differences between what people say and do, which is one of the most interesting areas of the privacy experience. Second, the research was done in the field of interest for this thesis, electronic commerce. Third, the research in question had quite a detailed description of the empirical research method and process that it was executed with. This provided a good base for replicating the research.

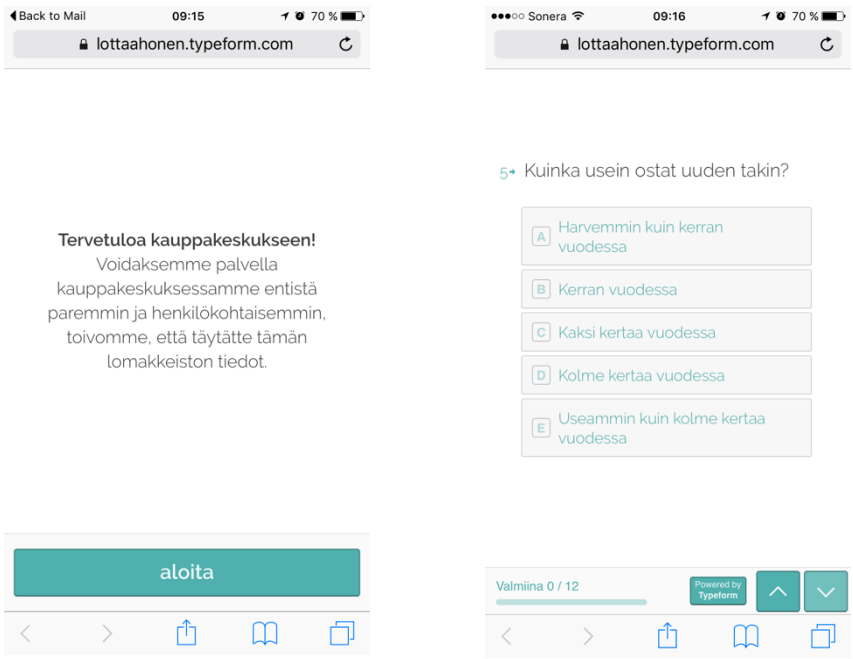

Figure 3-1: Screenshots of the prototype used in the research

For the purpose of this research, three different versions of the prototype were developed in order to see if the answers would vary in any way in different situations. This approach was adapted from previous research done by Culnan & Armstrong (1999) where they focused on looking into the propensity of people to give personal information in an e-commerce setting. In their research, they used two different scenarios: one where fair data collection practices were prominent and in the other they were not explicitly mentioned. In their research Culnan & Armstrong got different answers from these two different scenarios. The people who were more concerned about their privacy were more likely to not give their information when fair data collection practices were not prominent. In the other scenario, they found no difference between people who are concerned about their privacy and the people who are not. It was also in the interests of this thesis to see if there is a difference in what information

people give in different situations, in different states of awareness and control. The fundamental purpose of the different scenarios was to test how different levels of information transparency and control affects the experience.

The different versions of the applications had the same basic questions. The difference was made in showing the recipients of the information differently to the participants in three different versions. Version one did not have any indications whatsoever about who will receive the information. The second version displayed a list of the recipients at the end of the questionnaire, without any option to interact with the list. The third version had a list of the recipients of the information in the beginning and the user was able to control who would get his information.

The basic questions in the prototype consisted of 12 questions mapping out personal information needed in order to buy a jacket. These questions were based on a previous study made by Spiekermann et al. in 2001. In their study, Spiekermann et al. investigated drivers and hindrances of online interaction with the help of an experiment in an online store. This topic was very close to the first research question of this thesis. The research situation and questions in Spiekermann et al.'s research were suitable for answering the research question in this thesis. Thus, this previous study can be taken to be as an example in the research part.

In research done by Spiekermann et al., they tested their participants with a situation, where the participants were to buy either a jacket or a camera from an online store. In the store, there was a bot asking questions in order to help and guide the user towards what they are looking for. The researchers were interested in which questions the participants would answer. The questions were divided into four different categories: questions about the product itself, questions linked to the use of the product, questions that were personally related to the buyer but that also have an effect on the product recommendations and questions that were not linked to the product selection itself. For this thesis, some of the questions from the 2001 study were taken and modified to fit the scenario. Questions from each category were chosen in order to keep the same similar setting as in the original research. The questions were presented in the application starting with more general questions and ending with almost irrelevant questions, in order to see when the questions become too personal. The questions can be seen in Appendix B.

The initial hypothesis with this prototype was that the version three would be seen as the most pleasurable to use from the privacy perspective, thus creating the best privacy experience. This assumption was made based on the fact that it had the most transparent information about the recipients of the information as well as the option to choose to whom to actually give the information to. Version two was assumed to have the second-best privacy experience as it had transparency about the recipients. Version one was assumed to have the worst privacy experience. This hypothesis will

be measured based on how many answers does each version get compared to each other. The one with the most questions answered will be seen as having the best privacy experience.

As the questions were positioned in the prototype starting from the most general and becoming more personal and irrelevant, the second hypothesis was that people would stop answering to the questions at a certain point. The assumption was that the question "What do you do on your free time?" would be the question drawing the line.

3.2.2 Interview

The questions for the interview part were developed to function as conversation starters about how people view applications that collect personal information and how people usually act with them. They were held immediately after the tryout with the prototype in order to maintain a similar setting and situation. The interviews were semi-structured, where the questions were modified on the go to suit the certain situation. All of the interview questions followed pre-thought themes, that were the actual use of a personalized shopping application, why give or not give certain information, privacy perception of the prototype and general views on privacy. The questions included questions about the use of the application that the participants had just tried out as well as questions about their behavior usually in situations when personal information was asked. The list of pre-thought questions can be found in Appendix C.

3.2.3 Questionnaire

The questionnaire consisted of five questions, mainly following the main topics of the semi-structured interview. It had a four-point scale with options ranging from "I totally agree" to "I totally disagree". A four-point scale was chosen in order to get the participants to really think about their believes and to avoid too neutral responses.

The main purpose of the questionnaire was to support the empirical data with some numerical data. Thus, it was not based on any previously made questionnaire. It was made in order to ease the comparison of the three different versions of the prototype and to see more clearly which features had the most effect on the privacy experience. The questions can be found in Appendix D.

4. Results and analysis

This section presents the results and main findings of the two-folded empirical research: the internet questionnaire and the prototype experiment with the prototype, interview and questionnaire.

Based on the literature research, it can be noted that most of the analysis methods used in the field of privacy are numerical. For example, the study by Spiekermann et al. (2001), which is used as an example for creating the prototype used in this thesis uses regression analysis as an analysis method. They also used the PCIC index. These methods are suitable for big amounts of data. As the amounts of answers in the quantitative questionnaires were relatively small, statistically significant analysis was not possible to make. Thus, all of the results from the questionnaires were only approximates and should be treated as such. The interviews address a novel theme. The answers from the interviews were classified into different categories and analyzed first within these categories (Taylor-Powell & Renner 2003). The categories were formed by creating an affinity diagram based on the recorded notes from the interviews and taking the biggest clusters as the categories.

4.1 Internet questionnaire

The questionnaire was distributed in Finnish social media in December 2015 and it was available for about a month. The complete amount of answers was 23, with ages ranging from 20 to 67. 23 answers cannot be considered as a big amount of answers for an internet questionnaire but the results can still show some interesting trends.

The three most trusted entities were health care, banks and government, with Google and social media being the least trusted. The detailed results can be seen in Table 4-1. The reasons why people had chosen to trust these entities were quite unanimous. Seven out of the 14 respondents that answered to the question why they trust the entities stated in the first question, trust entities based on the image that these branches and entities are heavily regulated. It was also stated that people trust entities such as banks and insurance companies because they would be in deep trouble and harm their business if they misused or leaked private personal data.

"Banks would end up in huge problems if they would get caught leaking private data."

Health care	87 %
Banks	74 %
Government	65 %
Family and friends	48 %
Insurance companies	39 %
Colleagues	30 %
Public transportation	22 %
Google	13 %
Cloud services	13 %
Retail stores	4 %
Social media	4 %
Only myself	4 %
Nobody	4 %

Table 4-1: Who do you trust with your personal information?

People were more concerned about entities leaking their information to third-parties than on first-hand misuse. People were ok with sharing their information as long as it stays with that entity and creates value for the individual or the society. The same people seem to think that sharing personal information and proper use of it will make their lives easier while others see that giving away information will only make their lives more complicated.

> *"I think sharing my personal information with these will bring either personal or general benefit."*

What information people were ready to share was very divided. Three participants were willing to share only the most necessary information. 14 participants said that what they share depends on to whom they are sharing it to. They were ready to share quite much data under certain conditions. The general feeling was that the closer and personal the information recipients are, like family and friends, the more information they would be ready to share.

> *"Any relevant data that is needed to get better service IF the service provider (private companies) can guarantee that my data is removed when I stop using the service."*

When it came to the question about trusting one's competence in keeping personal information, over half of the respondents were confident in their own skills to safeguard their personal information (Table 4-2).

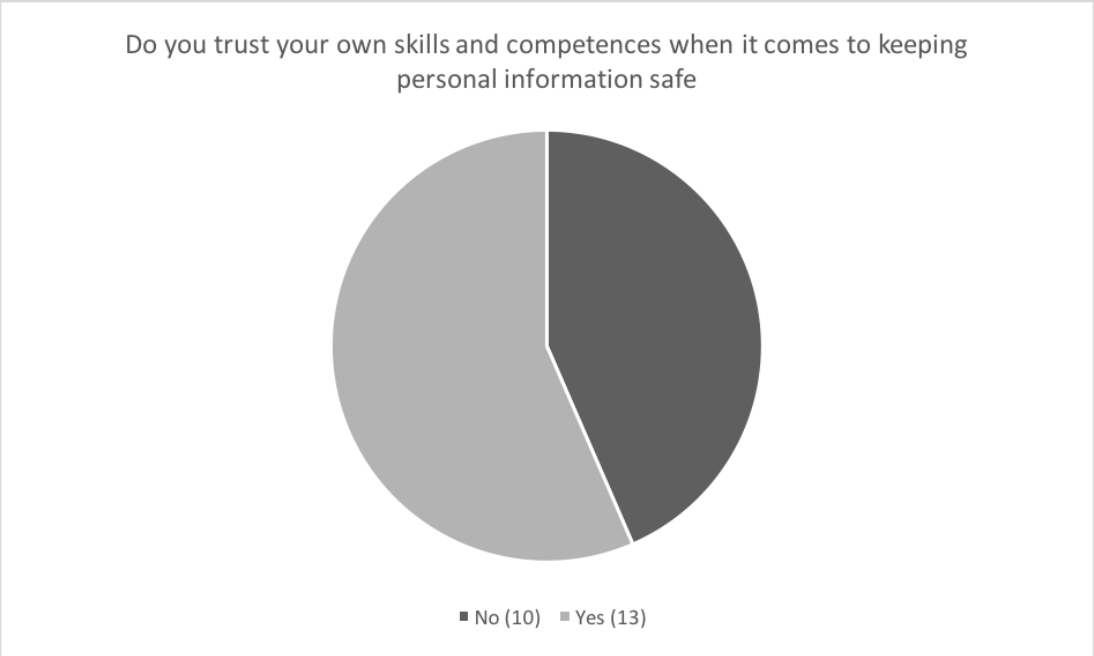

Table 4-2

According to the questionnaire conducted for this thesis, the three most trusted entities were heath care, banks and government. Retail stores and social media were found in the bottom of the list. The most trusted ones were entities that have their operations highly regulated by laws. This indicated that regulations seem to create an atmosphere of security and trust that reduces the negative privacy perceptions. This makes people think that these entities do not misuse personal information. Like one participant said, people trust these entities because misusing personal data would harm their operations and be against the law.

Based on the assumption that regulations create a feeling of security and trust, retail stores and social media are seen as entities with the least regulation as they are found at the end of the list.

4.2 Prototype with Interviews and questionnaire

All in all, 11 people were interviewed for this part of the research. The majority of the participants were in the age group of 25-35, as the purpose of the research was to target people who are already familiar with digital services and use them actively. One participant did not provide her age. For comparison, people from younger and older age groups were included. More women were chosen for the research than men but that was not intentional as gender was seen as an irrelevant factor in this study. The statistics of the demographic distribution of this research can be seen in Table 4-3.

Age group:	15-24	25-35	35-55	N/A		Gender:	Male	Female	Other
	2	6	2	1			3	7	1

Table 4-3: Demographics of the participants

Three of the participants got to test version one of the prototype, without any indication about information recipients. Five participants tried out version 2, the one with a list of recipients at the end of the questionnaire. Three participants tested version three, the prototype with the possibility to choose the recipients of the information.

The participants were encouraged to speak as they tested the prototype. They were also interviewed immediately after using the prototype. The questionnaire was given to the participants as they were being interviewed.

4.2.1 Prototype

The given scenario was understood well and the participants were answering according to it. In general, people were quite open and positive while answering the questions in the prototype. Still, six out of eleven people did leave one or more question unanswered.

Version 1

Prototype version 1, without the list of recipients, was tested by three participants who all happened to be female. They answered to 86 per cent of the questions. One of the participants answered to all of the questions, while one left four questions unanswered. The last participant left one question unanswered. All of the unanswered questions were different questions.

Version 2

Prototype version 2, with the unmodifiable list of recipients, was tested by five participants. All in all, they only left three questions unanswered, having their response rate at 95. Two of the participants answered to all of the questions and three participants left one question unanswered. All of the unanswered questions were different questions.

Version 3

Prototype version 3, with the modifiable list of recipients, was tested by three participants. In total, they left five questions unanswered. Their response rate was 86 per cent. One participant left three questions unanswered while the two others left one question each unanswered. Two of the participants did not answer to the same question.

The differences in versions 2 and three can be seen in figure 4-1.

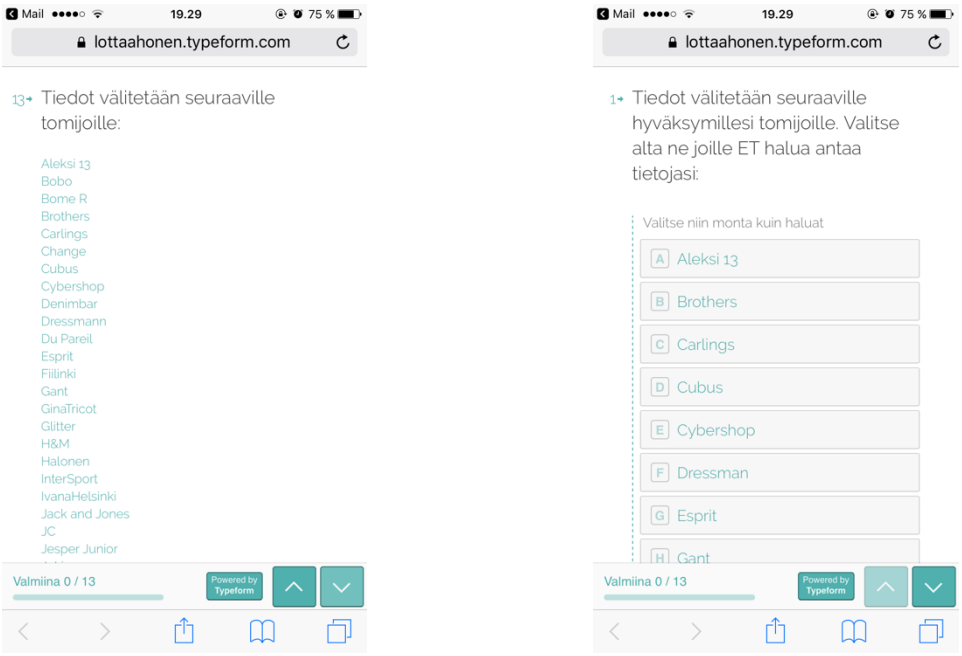

Figure 4-1: Differemces between prototype versions 2 and 3. On the left version 2 with an informative list of recipients. On the right, version 3 with a selection list.

The results from the prototypes were analyzed by comparing the answers from the different versions to each other. The original study by Culnan & Armstrong (1999), where the method of having multiple version was adapted from, uses a discriminant analysis method as their analysis method. This method was not suitable for this thesis as there were only 11 answers and this would not trustworthy answers with the original analysis method. Thus, the answers from the three different prototype versions will be compared and analyzed qualitatively.

Two hypothesis were created for this experiment. The **first initial hypothesis** with the prototype was that the version three, the prototype with the option to choose to whom to give the personal information, would be considered as the best when it comes to perceived privacy. The first version was thought to make people question the missing information about the recipients. The second version was thought to be off-putting as it shows the list of recipients but does not give any control over the list.

The first hypothesis was based on the literature research, where transparency and choice had been discovered to be closely linked with creating a desired privacy experience (Section 2.5). Transparency was said to create ease the anxiety of the unknown (Oulasvirta et al. 2014). The freedom of choice was said to empower the user and make the experience more pleasurable (Culnan & Armstrong 1999).

Based on the results, it seems that the first hypothesis did not hold true. The best response rate, 95 per cent, was recorded for the prototype version two. Prototype versions one and three got the same response rates of 86 per cent. It seemed that the version

The **second hypothesis** was that people would consider the question "What do you do on your free time?" as intrusive enough and where the participants would stop answering. This hypothesis came from the interests of the industry as they were keen on looking into finding the possible creepy line where people feel too uncomfortable to answer questions.

The second hypothesis held partially true. As it could be seen from the results, the question "What do you do on your free time?" was left unanswered the most. But the participants did not stop answering at this question. All of the participants that left the question unanswered did answer the last two questions that came after it. It seems that the question was intrusive enough but the assumption that people would stop answering altogether after it could not be observed in the research in this thesis.

The overall response rate was 89 per cent. In total, the most unanswered question was "What do you do on your free time?" having four out of eleven people not answering it. When looking at the response rates of the different question categories from all of the version combined (Table 4-4), we could see that the category personal questions independent of the product had the lowest response rate of 77 per cent. This category included the questions "How often do you buy a jacket?" and "What do you do on your free time?"

	Product attribute questions	Usage oriented questions	Personal questions supporting product selection	Personal questions independent of product selection
Response rate:	91 %	86 %	97 %	77 %

Table 4-4 : Response rates of different question categories

The participants seemed to be most willing to answer personal questions that helped in selecting the right product. Next came the questions about the actual product and questions about the usage of the product. Clearly the most irrelevant questions related to the action of buying a jacket were left unanswered most frequently.

4.2.2 Interviews

Because the interviews were made in the same situations as the prototype tests, they were recorded as handwritten notes by the interviewer. After each interview, the notes were written out in a digital format. After completing all of the interviews, the answers were analyzed by creating affinity diagrams of them. This revealed six different clusters in the answers (Figure 4-1).

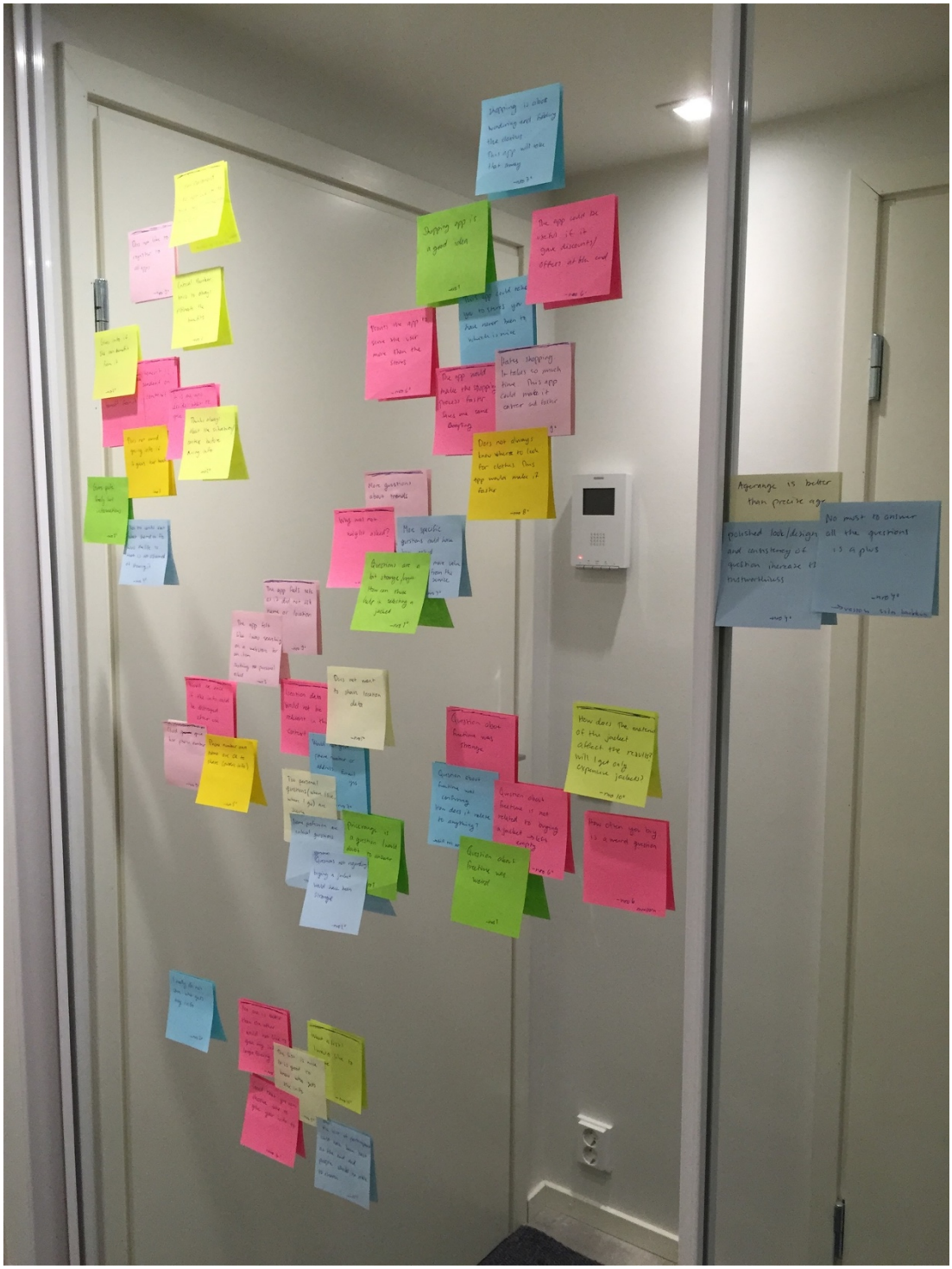

Figure 4-2: Clusters emerged from interviews

The clusters were:

- The usefulness of the application (12 comments from 11 different participants)
- Comments on the weird questions in the application (6 comments from 5 different participants)
- Comments about the list in the different versions (6 comments from 6 different participants)
- What type of information can be shared and what not (12 comments from 9 different participants)
- What creates a good privacy experience (in this application) (3 answers from 2 different participants)
- Views on privacy in general (11 comments from 7 different participants)

The usefulness of the application

Five out of 11 participants thought that the idea behind the personalized shopping application was good. Especially people who do not like shopping that much or do not have time to do it, seemed to like the idea of an application that would save them time from visiting different stores.

> *"I hate shopping, especially when I have my kids with me. This application would make it so much more efficient." – Interviewee no. 9, woman, 35-55 years*

Three participants did not find the application suitable for them. It became clear that people who enjoy shopping and have the time to do it, would probably not use the application as it takes away the enjoyment of wandering around stores.

> *"To me, shopping is enjoyable. I like to wander through stores and feel and see different clothes. It might turn it to only a boring execution of going in and out of the store with this application." –Interviewee no. 7, woman, 15-24 years*

Three participants saw the application as a positive thing, but were not completely sure if they would use it, at least in its current form. They were wondering if the application could actually recommend anything based on the current questions. Two of the participants felt that there were some essential questions missing. For example, weight was one question that was missed, as it would affect the selection of the jacket.

> *"There could have been even more specific questions about the jacket, such as how personal should the style be. That would bring more value to the service." – Interviewee no. 4, man, 25-34 years*

All in all, it seems that there could be a need for this type of service amongst people who are too busy to wander through stores. A personalized shopping application could guide people faster to the right stores and thus speed up the shopping.

Comments on the questions in the application

The necessity of some of the questions were challenged. Four participants wondered aloud about the question "What do you do on your free time?". The participants found it irrelevant to the process of buying a jacket and that was confusing.

> *"I left the question about free time empty as it has nothing to do with buying a jacket. I don't want to give too much information." – Interviewee no.6, woman, 25-34 years*

Two out of the four participants that doubted the question concerning the usage of free time did answer the question. One of them explained that he anyway answered because he suspected that it might tell something useful about the person answering the questions regarding buying a jacket.

> *"The question about free time was a bit suspicious, but maybe it tells something about the personality of the person." – Interviewee no. 1, man, 25-34 years*

The questions about the material of the jacket and "How often do you buy a jacket?" raised some questions. The participant doubting the question "How important is the material of the jacket to you?" suspected that by answering this question the prototyped service would only suggest expensive options. However, she answered the question. The participants doubting the question "How often do you buy a jacket?" also answered it.

The majority, four out of six of the participant who criticized some of the questions during the interview had still answered to them while trying out the prototype. There could be multiple reasons why they had decided to act differently than they were thinking. One could be that the participants have actually seen the prototyped service to bring value to them.

Comments about the list in the different versions

Participants that tried the version one, without any recipient list, did not mention that they would be missing the information about who their personal information is going to. Participants who tested the versions two and three with the recipient lists felt generally good about the lists. Two of them were also surprised when they got to see who exactly gets their information. Participants with the selectable list (version 3) appreciated the fact that they could see where their information was going to be used and especially that they had a chance to have an effect on the recipients.

> *"The list was nice. It is good to see where my information is going." – Interviewee no.5, woman, 25-34 years*

Participants who tested the prototype with the informative list without any selections (version 2) were also pleased that they could see where their information was going, but they seemed to think that the list was a bit strange and that it could use some improvements. They wanted the opportunity to modify the list, to choose who to give

their information to. Participants with the version without any type of list seemed to not think about or care about where their information is going. None of the participants expressed their wish to see where their data is going.

> *"Oh, my information goes to all of them! It would have been fair to tell that in advance. It would have been nice to get to choose who gets my personal data." – Interviewee no. 4, man, 25-34 years*

Two out of the three participants that tested the version three did limit the recipient list. Their reasons for excluding some stores on the list were that the participants did not want to give their information to large corporates. The female participant also explained that she wanted to exclude all of the stores that sell only men's clothes and stores she had not visited ever.

All in all, it seems that the list is a feature that is appreciated as it reveals the recipients. Nevertheless, the participants without a list did not yearn for one. Could it be that by not showing the recipients, the participants did not know to miss it?

What type of information can be shared and what not

When asked about what type of information participants would not be willing to give in this scenario, two common themes were found in their answers. Half of the participants were strongly against sharing their location data. One participant was not willing to give out the information of the area she lived in. Even though knowing one's location could make the suggestions from the application more accurate by weighing more the store nearby, six out of eleven participants mentioned that they would not feel comfortable sharing their precise location. Other sensitive types of information were name, occupation, exact age and other very personal and identifiable information.

The other common theme was that the participants were willing to share more information if they were to gain something by revealing it. The participants thought through every question and weighed how important the question is for the application to give them the best possible value.

> *"I always think about the situation and context when giving personal information to a new service. I give it If I feel that I can benefit somehow from it." – Interviewee no. 2 woman, 15-24 years*

One participant was pleased that there was no question about what price range the jacket should be in, as the participant assumed that based on the answer the application could predict that he was a poor student. Any information that could identify the participant easily and was not necessary for the service were on the list of information not to share. Some exceptions were also found. Phone number was seen okay to give by two of the participants whereas three of the participants were heavily against sharing their phone number. Email was seen as acceptable to share as it is nowadays required

in so many places. One participant was willing to share her name, address and phone number as they are public information that are freely available.

"I could see myself giving my number if I would get beneficial add through WhatsApp." –Interviewee no.9, woman, 35-55 years

On the other hand, some of the participants seemed to be quite open about themselves to the public. One participant explained that he lives his life in a way that he dares to publish to the world. He shares a lot about his life but is aware about the consequences and thus is in control of what personal information gets out to the public.

What creates a good privacy experience (in this application)

The participants were not able to specify in detail what makes a good privacy experience. A couple of theme were able to pinpoint a few factors that they believe are good design for privacy in the prototype in question. They mentioned that having many questions out of the context make people question the reliability of the prototype affecting the privacy experience. Having too specific and personal questions raised alarms as they were seen as easy ways to personify an individual.

"A well-designed application with coherent questions increase the trustworthiness of an application" –Interviewee no.4 man, 25-34 years

Other factors that were mentioned to enhance the privacy experience were the possibility to not answer all of the questions, seeing and having a say on who gets the information and the ease of use. These create a feeling of voluntary information collection, a process that is in control of the user. In addition, the ability to destroy all personal data was mentioned by two participants as a feature they would appreciate.

Views on privacy in general

From the affinity diagram, it could be seen that people could be divided into three different groups based on their preferences of sharing personal information: participant very particular about their personal information, participants very liberal about what they share and participants between these two ends.

According to the interview data, one of the participants was very particular about giving out personal information and privacy in general. She described herself as a bit paranoid about what retailers know about her. She felt that it would be better if people could voluntarily provide information to retailers rather than the retailers spying on people's shopping behavior. This one participant was the only one that could be clearly separated to this group.

"I am paranoid. I do not want to give my information to the stores." – Interviewee no. 10 woman, 25-34 years

Two of the participants were describing themselves as quite relaxed about what they share in the public. These participants were very open to the public, especially different social media channels, about their daily lives. Even though they were very liberal about sharing personal information, they were also quite rigorous about where they shared their information. They were not sharing everything everywhere but kept somehow track of where they post and share information.

"I want to live my life so that I'm not ashamed to show it to the public" – Interviewee no. 4 male, 25-34 years

According to the affinity diagram, five participants could be placed somewhere in between these two previous groups. They did not feel that comfortable about data collection and were cautious about what they were sharing but were not totally against it. The common trait that these participants had was that they contemplated heavily on the possible benefits in different contexts, thus they were quite picky about the information that they were sharing.

"I give information if I see it relevant for the service to work. I always try to assess the benefits before giving personal information." –Interviewee no. 2, woman, 15-24 years

Three of the participants were not able to describe their general views towards privacy.

Based on these results, it was possible to divide the participants into four different categories based on their general views on privacy. The participants were categorized as privacy unconcerned, privacy concerned, very privacy concerned or privacy undecided.

4.2.3 Questionnaire

All of the participants answered all of the questions in the final questionnaire. The total results of all the different versions can be seen in Table 4-5, where answers from different versions are presented in parenthesis in the order: version 1, 2 and 3.
The participants seemed to mostly agree with the statements presented to them. Nobody fully disagreed with any of the statements.

	Fully agree	Agree	Disagree	Fully disagree
I would be ready to use the personalized shopping app	**4** (1,3,0)	**5** (2,0,3)	**2** (0,2,0)	**0** (0,0,0)
I felt that the application was safe	**3** (1,2,0)	**6** (2,2,2)	**2** (0,1,1)	**0** (0,0,0)
I feel that applications, such as the one just tested, do not violate my privacy	**2** (1,1,0)	**5** (0,3,2)	**4** (2,1,1)	**0** (0,0,0)
I am willing to give my personal information to companies if I get something in return	**2** (1,1,0)	**7** (2,3,2)	**2** (0,1,1)	**0** (0,0,0)
I am generally worried about my privacy	**1** (1,0,0)	**5** (1,1,3)	**5** (1,4,0)	**0** (0,0,0)

Table 4-5: Total answers of post-prototype questionnaires. In parentheses results from different versions.

As the amount of participant in total and for each different prototype version were so small, no reliable quantitative analysis can be made on the results. Thus, the results will be analyzed qualitatively one statement at the time.

The first statement in the questionnaire was "I would be ready to use the personalized shopping app". Participants with the version one of the prototype (without a list) agreed (2) or fully agreed (1) with this statement. The answers from participants with the version two (with a non-selectable list) varied more. Three of them fully agreed while two of them disagreed with the statements. The participants with the version three (with a selection list) were unanimous. All of them (3) agreed with the statement. Based on the results it seems that participants with version two are the most skeptical ones about using a personalized shopping application.

When comparing what type of people have answered how, one thing becomes clear. The ones that have answered most positively to the first question were also the ones that had said during the interview that they see the personalized shopping application as something they would use. This was regardless of what version the participants have tested. The ones that disagreed with the statement were participants who either loved traditional shopping or were very skeptical about handing out any personal information.

The second statement was "I felt that the application was safe". Participants testing version one agreed (2) or fully agreed (1) with this statement. There was again a bit more dispersion among the participants testing version two. Two of them totally agreed, two agreed and one disagreed with the statement. Participants testing version three agreed (2) and disagreed (1) with the statement.

No clear common factor could be found between the participants who fully agreed or disagreed with this statement. People who considered themselves as quite liberal when it comes to sharing personal information were found in both ends of the answers.

"I feel that applications, such as the one just tested, do not violate my privacy" was statement number three. Participants with version one fully agreed (1) or disagreed (2) with the statement. Participants with version two fully agreed (1), agreed (3) and disagreed (1) with the statement. With the third version, participants agreed (2) or disagreed (1).

The last two statements were questions that really were not related to the prototype or its version. Thus, these is not that much interest in analyzing the answers divided by the different versions.

The fourth statement was "I am willing to give my personal information to companies if I get something in return". Two participants totally agreed, seven agreed and two disagreed. The participants seem to be generally positive about sharing their information if there is something that they get in return and that they feel that it is valuable.

The fifth and the final statement was "I am generally worried about my privacy". One participant fully agreed, five agreed and five disagreed with this statement. Based on this it could be said that the participants are pretty equally divided into privacy concerned and unconcerned. Only one participant seemed to be very concerned about her privacy. Based on this, the participants could be divided into three different groups based on their views on privacy: privacy unconcerned, privacy concerned and very privacy concerned.

The question about general views on privacy was asked also in the interviews. There seemed to be a difference between the participants' answers between the interview and the questionnaire (Table 4-6). Only five out of 11 participants answered similarly in both even though the interview and questionnaire were basically conducted one after the other.

Participant	**Interview**	**Questionnaire**
1	unconcerned	unconcerned
2	concerned	unconcerned
3	undecided	very concerned
4	unconcerned	unconcerned
5	concerned	concerned
6	undecided	concerned
7	undecided	unconcerned
8	concerned	concerned
9	concerned	concerned
10	very concerned	concerned
11	concerned	concerned

Table 4-6: Differences in privacy segments between interview and questionnaire

The categories from the different methods are very similar. It could be said that altogether four segments were identified based on these two different methods: privacy unconcerned, privacy concerned, very privacy concerned and undecided.

5. Discussion

This section discusses the major findings found in the research. These findings will be analyzed and their meaning and importance will be discussed.

5.1 Who do people trust with their information?

In this section, the results from the internet questionnaire asking with whom do people trust with their personal information will be analyzed and compared with two similar questionnaires carried out quite recently.

Morey et al. (2015) have conducted a similar worldwide internet survey about with whom would people trust with their personal information. In their research, social media firms, such as Facebook, were rated as the least trustworthy and the most trustworthy were doctors and credit card companies. See the all of the results in Table 5-1.

Primary care doctors	87 %
Payment or credit card companies	85 %
E-commerce firms	80 %
Consumer electronics firms	77 %
insurance companies	76 %
Banks	76 %
Telecom carriers	73 %
Technology firms	70 %
Internet giants (Google, Yahoo, etc.)	68 %
Governments	66 %
Media and entertainment companies	61 %
Social media firms	56 %

Table 5-1: Categories that were seen trustworthy or completely trustworthy (Morey et al. 2015)

This is quite well aligned with the results gotten from the internet questionnaire conducted for this thesis. In this research, social media was in the bottom four alongside with the following options: retail stores, only myself and nobody. Also, the options selected as the most trustworthy are quite in line with each other. Both researches have healthcare on the top.

The biggest difference in these researches seemed to be the position of government. In Morey et al's research, government was ranked in the lowest third whereas in the research for this paper, government was in the top three. This could be explained with the different demographics of the two researches. The questionnaire done for this thesis was answered by mainly Finnish people, whereas Morey et al's questionnaire was distributed worldwide.

Fleming & Yu (2015) conducted a similar survey in The United States (Table 5-2). Their research indicated towards similar findings regarding the least trustworthy options. Social networking websites or applications were the least trustworthy by far. In the bottom three were also the federal and state governments. Social media being least trusted was a common thread between all of the three different surveys.

Some differences were found in who people trust the most. Fleming & Yu's top three most trustworthy consisted of banks, companies you regularly do business with and brick-and-mortar retailers. They have separated physical stores and online retailers to two different entities. Interestingly, these two entities have very different placings in the results. Furthermore, the results from the questionnaire for this thesis as far as retail is concerned, were on common ground with the results for the online retailers in Fleming & Yu's research. It could be that the people that have answered the questionnaire for this thesis have interpreted retail stores as online stores or that people are not that happy with the multiple loyalty clubs and cards that many retail stores in Finland have.

Your primary bank	91 %
All companies you regularly do business with	78 %
Brick-and-mortar retailers	69 %
Health insurance companies	68 %
Credit card companies	66 %
Your email provider	63 %
Your cellphone platform (iPhone, Android or Windows)	63 %
Your cellphone carrier	59 %
Online retailers	58 %
Your state government	52 %
The federal government	45 %
Social networking websites or applications	23 %

Table 5-2: Categories that were seen trustworthy or completely trustworthy (Fleming & Yu 2015)

Yet again, a big difference could be found in the positioning of the government. Similarly, as in Morey et al.'s research, Fleming & Yu's research also showed the government in the bottom of the list. The differences between trust in government between Finland and the rest of the world is an interesting subject.

All in all, it seemed that social media was the least trusted, no matter where the research was done. For the most trusted ones there was not that clear pattern or common reasoning. Banks and the health care industry seemed to be generally seen as worth of trusting whereas the placing of the government varied. The reasoning in the questionnaire for this thesis for having heath care and banks as the most trustworthy, was that those industries are heavily regulated. Therefore, they must obey the rules

and regulations set for them. This makes people feel safe. Interestingly, in Morey et al.'s and Fleming & Yu's researches, the reasons for the selections of the most trustworthy entities were not mentioned.

5.2 What type of data were people willing to share?

Previous research showed that different types of data have different sensitivity levels when it comes to sharing personal information. The most sensitive types of data were financial and health records. Medium sensitive types of data were media usage and location. The least sensitive types were gender, age and other similar data. (Horne & Horne 1998; Phelps et al. 2000; Malheiros et al. 2013; Rose et al. 2012) In this research, the participants were not willing to give their location, name, occupation, exact age or phone number. According to previous research, all of these, except location, belong to the group of least sensitive data.

Based on the internet questionnaire done for this thesis and similar ones before, it can be said that the receiver of the personal information also has an effect on what type of data people are willing to share. The more trusted the receiver is, the more easily are people willing to share personal information. (Morey et al. 2015; Fleming & Yu 2015)

From the interviews, it became clear that the participants were willing to give any personal information that could benefit them. The perceived benefit is something that the participants assess individually and it might vary quite a lot between people and situation. The combination of retail not being the most trusted when it came to giving personal information and people not seeing sharing location as something they could benefit from could explain why people were reluctant in giving their location information.

This thesis concludes that what information people are willing to share depends on who the receiver is and on the perceived benefit that the individual feels that he or she gets in return of sharing that information. Thus, the question of what type of information are people willing to share is dependent on the context.

5.3 Hypothesis

For the prototype tests two hypotheses were created. Hypothesis one assumed that prototype version three would be seen as the best from the privacy point of view. It was expected that this version would get the most answers from the participants. Hypothesis two assumed that people would find the question "What do you do on your free time?" too personal and out of the context that they would not answer this question or any of the questions after this. It was assumed that this question would create the creepy line.

The first hypothesis was proven false in the tests. The prototype version number two

had the highest response rate of all the different versions, whereas versions one and three had lower response rates that were identical with each other. This is contradictory with the findings in the literature review. Literature stated that more control results in increased trust, which again encourages more answers (Brandimarte et al. 2012; Acquisti et al. 2015). This did not hold true in this research. No clear reason for this could be found from the data collected for this research. The participants' views on privacy were mixed throughout the different prototype versions.

The second hypothesis was proven partially incorrect. The question "What do you do on your free time?" was left unanswered the most but people still answered the questions about height and gender that came after this one. When asked why the participants did not answer this question, the most common answer was that participants felt that it had nothing to do with trying to buy a jacket. They could not see that the application would need that answer in order to give them good suggestions about new jackets. They felt that it would not bring any extra value to them if they answered the question.

This finding could be explained with the phenomenon of privacy calculus. Individuals usually tend to make assessments about how their personal information will be used and how that usage will be reflected back to the individual (Culnan & Armstrong 1999). With the prototype in this thesis, the participants have made an assessment that the question "What do you do on your free time?" is not relevant for the application in order to get most benefits out of the application. However, the two following questions about age and gender had been seen crucial information for the selection of a jacket. This was confirmed in the interviews.

5.4 Factors of privacy experience model

All of the factors that were presented as a part of the privacy experience model in the section 2, were also mentioned in one way or the other in the empirical research by the participants. Additionally, there was one factor that arose clearly from the - empirical research but was not mentioned in the privacy experience model. This factor could be called perceived benefit. Many participants stressed the importance of getting something in return for giving out personal information. If they did not feel that they gain anything by giving some particular information, they would simply not give that information.

In the literature research conducted for this thesis, there were a couple mentions about a phenomenon called privacy calculus, that could explain this. Individuals' have the ability to assess possible risks and benefits that might be caused by disclosing personal information. This is called privacy calculus. (Culnan & Armstrong 1999) Especially when an individual is taking part in electronic commerce transactions, they make calculations about if the benefit they are about to gain is worth the possible risk associated with collecting personal information (Gurung et al. 2014). Individuals might

do privacy calculus on single transactions or about the companies that they transact with (Gurung et al. 2014).

Privacy calculus was not taken into the privacy experience model, as it was not found that many times in the literature. However, based on the empirical research, privacy calculus should be considered as a worthy factor when thinking about privacy experience. It could be added to the individual factors side of the model. The updated privacy experience model can be seen in Figure 5-1.

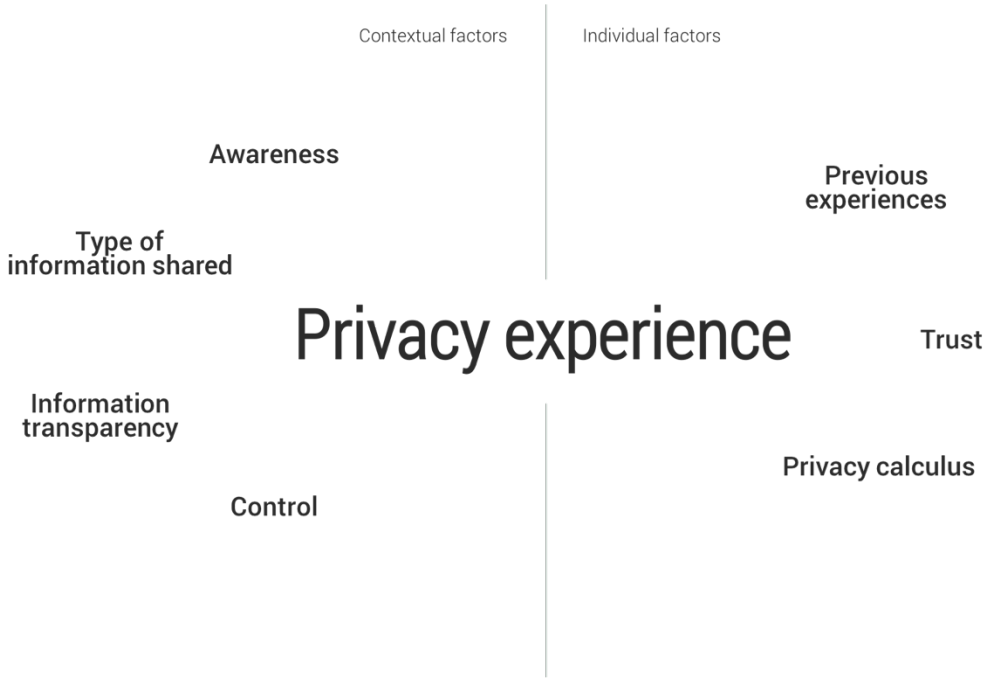

Figure 5-1: The updated privacy model

5.5 Westin's privacy segments

Based on the interviews and the questionnaire, the participants could be divided into four different groups based on their views on privacy. The groups were privacy unconcerned, privacy concerned, very privacy concerned and undecided. These emerged groups could be compared to Westin's (2003) privacy segments. Westin has defined three different privacy segment groups: privacy fundamentalists, privacy pragmatists and privacy unconcerned. How do these groups compare?

Westin's group privacy fundamentalists were described as people who are extremely protective of their privacy. They were the people who are proactive in refusing to give personal information to companies. (Krane et al. 2002) This description is similar to the one in this research. The group very privacy concerned consisted of a person that described herself as "a bit paranoid" when it comes to information collection. She did not trust companies to keep her information safe and would prefer voluntary data collection.

Westin's group privacy unconcerned were described as people who feel that they benefit from sharing personal information. They are not that strict on what they share. (Krane et al. 2002) The group privacy unconcerned in this thesis were described as people who share a lot of information as they see that they benefit from it. However, they still are aware of what they share and where.

Westin's last group, privacy pragmatists, were described as people who think about the benefits of handing out personal information. They make their decisions based on if they see that they benefit from it in any way. (Krane et al. 2002) The equivalent group from this thesis, privacy concerned, were described as people who contemplate heavily on the pros and cons of sharing personal information and are very selective when it comes to data sharing.

This thesis identified one additional segment on top of the before mentioned three segments, undecided. They were people who could not exactly say what they thought about privacy. This segment is not included in Westin's segments. Other than that, the segments seem to match quite well with each other. It can be said that Westin's privacy segments were identified from the empirical research. However, not all participants identified themselves in the same privacy segment in the interviews and the questionnaire. Interestingly, only five out of 11 participants could be categorized in a similar group in both methods.

The before mentioned findings go along well with the current views in research. The dominant perspective nowadays is that peoples' views on privacy change according to situation and context. Thus, it is not possible to say that a person belongs permanently to one privacy segment (Consolvo et al. 2005; King & Hoofnagle 2008; Malheiros et al. 2013).

5.6 Privacy paradox

This research in this thesis indicates similar results as before, when it comes to the differences in peoples' attitudes and behaviors. Based on the results from the prototype and the interviews, it can be stated that people tend to have a difference between what they say about their attitudes towards privacy and how they actually act when it comes to giving personal information. For example, participants criticized some of the questions as being out of the context but they still answered them. Also, the participants who described them as being very concerned about privacy in general answered all of the questions in the prototype.

This phenomenon is called privacy paradox. Privacy paradox has been described as the difference between individuals' privacy-related attitudes and their actual behaviors regarding privacy protection and information disclosure (Norberg et al. 2007). In this thesis, the difference could possibly be explained with the notion that many

participants did feel that the application could bring more value to their lives and they saw that the questions supported the aim of the prototype. People tend to be more willing to share their information if they saw that they could benefit from it. People also have the tendency of thinking that nothing bad will happen to me. More precisely, in this case of giving personal information, people feel that they will be only one grain of information in the pool of data.

The setting in the empirical research, having a personalized shopping application, could be compared to store loyalty cards. Both collect personal information about what people like to wear and buy. They both give discounts and suggestions back in exchange. In previous research regarding privacy behavior and loyalty cards, people were concerned about how the data collected by the loyalty cards would be used. However, people did not take any actions in order to protect their data. They saw that the benefits were greater than the risks. This is a good example of privacy paradox. (Kang et al. 2015)

6. Conclusions and recommendations

This section discusses the meaning of the results on a more abstract level and gives recommendations for which way research should head in the field of privacy experience.

6.1 Answers to the research questions

The main research question was: **What type of data are people willing to share about themselves in a retail context?** In general, people are willing to give information that is seen as crucial for the service to work. But if they are asked to give some information that they feel is out of the context, they are not willing to give that data. What type of data people are willing to give was also seen to be dependent on what was the benefit that the individuals felt they could get in return? People are starting to expect something in exchange for their personal information.

The first supporting research question was: **What is privacy experience?** The base of the answer to this question was established in the literature research, where six of the main factors affecting privacy experience were identified. The model was then complemented with one factor that emerged in the interviews. The complete privacy experience model formulated in this thesis can be seen in Figure 6-1.

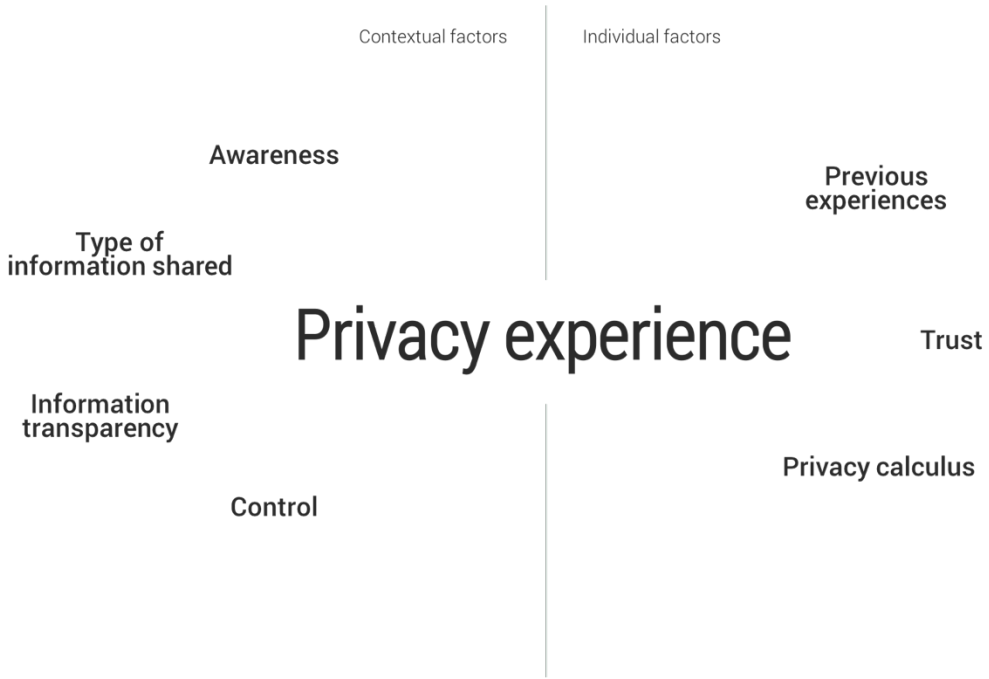

Figure 6-1: Model of the privacy experience

The second supporting research question was: **How can privacy experience be measured?** In previous studies, mostly only qualitative methods were used to measure different factors of the privacy experience. The most common ones were Concern For Information Privacy model (CFIP), Internet Users' Information Privacy Concerns scale (IUIPC), Private Consumer Information Cost (PCIC) and Privacy Concern Scale (PCS). All of these were based on Westin's privacy segments and did not include qualitative methods.

All of the research questions were aiming at answering the objective of this thesis. The objective was to research **how much personal information are people willing to share about themselves, especially in retail context**. Based on the literature review and the conducted research, it can be said that people are willing to give personal information they see relevant for the service to use. They will be more willing to give personal information if all the aspects of privacy experience (Figure 6-1) are taken into consideration. Thus, there is no clear, unambiguous way to answer this question. It is always dependent on the situation and context.

6.2 Limitations

This research was done on a fairly new subject stemming from the industry. Thus, there were some restrictions, especially related to conducting the empirical research part. The aim of the empirical research was to do multiple small and quick interviews in order to be efficient. Due to changes and limitations in the schedule, not so many interviews were carried out in the end. More interviews would have given more credibility for the results. The low number of interviews have an effect on how well the results can be generalized to a larger sample.

The look-and-feel of the application was not as initially planned, due to time limitations. An alternative plan had to be used, as it turned out that a custom-built application would not be possible to build in the wanted timeframe. Thus, an alternative solution was taken into use. The participants were aware that the prototype used was on a concept level. The participants were advised before the test to treat the situation as it would be real. Despite this, the participants might have thought that it is safe to answer all of the questions as the research was done with a prototype and not with a real application. A customized application, with the right look-and-feel might have given a different kind of impression about the situation and could have affected the answers.

For the empirical research, only around ten questions were made available from the original research by Spiekermann et al.(2001). If more questions would have been found, the questions for this research could have been designed more in detail and according to the original research. Now the ten found questions were just modified to fit the situation in the current research.

6.3 Theoretical implications

The privacy experience model created in this thesis is one of the first attempts to define the term privacy experience. As there was no clear definition of privacy experience before this thesis, this new model gives a basis that future research can build upon. This model should be tested in future research in order to validate it further. As the model is mostly based on previous literature, both theoretical and empirical research should be conducted.

6.4 Implications for the industry

The findings in this thesis should be taken into consideration when developing future services. The reason why people do not give their personal information was that they do not feel that they get anything in return or that that piece of information is not crucial for the service to work. Thus, when developing new services that utilize personal information from the users, one needs to think what information the service actually needs. For example, if there is a photo editing application being developed, does it need to know the user's contacts.

There is no such thing as a clear creepy line where people would always stop and think. It is heavily dependent on the user, the situation and the service provider. No single step-by-step guide can be given on how to develop services with good privacy experience. Based on this thesis, a good starting point is to take the factors in the privacy experience model into consideration when starting development.

The results of this thesis, especially the privacy experience model, has already been used in some projects in the industry. These projects have been related to the new EU regulation General Data Protection Regulation GDPR and how to manage one's personal information.

6.5 Future work

In the future, it would be interesting to get more qualitative research with bigger sample sizes on the subject of privacy experience and the different factors affecting it. Qualitative research could reveal new paths and revelation in some of the burning questions inside the field. It could bring some deeper insight to the question of what is the absolute reason behind privacy paradox, why do people say one thing and still act the opposite.

Another interesting topic to continue research on would be privacy segments and their applications in the modern information society. One question for future research could be that what causes people to change from one privacy segment to another. What is the dominant factor that causes a shift from one segment to another?

7. References

Ackerman, M.S., Cranor, L.F. & Reagle, J., 1999. Privacy in e-commerce. *Proceedings of the 1st ACM conference on Electronic commerce - EC '99*, pp.1–8. Available at: http://dl.acm.org/citation.cfm?id=336992.336995.

Acquisti, A., Brandimarte, L. & Loewenstein, G., 2015. Age of Information,. *Science*, 347(6221), pp.1–4.

Acquisti, A. & Grossklags, J., 2005. Privacy and Rationality in Individual Decision Making. *IEEE Security and Privacy*, 1, pp.26–33.

Allen, M.W. et al., 2007. Workplace surveillance and managing privacy boundaries. *Management Communication Quarterly*, 21(2), pp.172–200. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2009-10403-006&site=ehost-live.

Awad, N. & Krishnan, M., 2006. The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS quarterly*, 30(1), pp.13–28. Available at: http://www.jstor.org/stable/25148715.

Bélanger, F. & Crossler, R.E., 2011. Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems. *MIS Quarterly*, 35(4), pp.1017–1041.

Blackmer, W.S. (2016). *"GDPR: Getting Ready for the New EU General Data Protection Regulation".* Information Law Group. InfoLawGroup LLP. Available at http://www.infolawgroup.com/2016/05/articles/gdpr/gdpr-getting-ready-for-the-new-eu-general-data-protection-regulation/

Brandimarte, L., Acquisti, a. & Loewenstein, G., 2012. Misplaced Confidences: Privacy and the Control Paradox. *Social Psychological and Personality Science*, 4(3), pp.340–347.

Buchanan, T. et al., 2007. Development of Measures of Online Privacy Concern and Protection for Use on the Internet. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY*, 58(2), pp.157–165.

Cavoukian, A., 2009. Privacy by Design The 7 Foundational Principles Implementation and Mapping of Fair Information Practices. *Information and Privacy Commissioner of Ontario, Canada*.

Chellappa, R.K., 2002. Consumers ' Trust in Electronic Commerce Transactions : The Role of Perceived Privacy and Perceived Security. , pp.1–38.

Chellappa, R.K. & Sin, R.G., 2005. Personalization versus privacy: An empirical examination of the online consumer's dilemma. *Information Technology and Management*, 6(2–3), pp.181–202.

Chen, K. & Rea, A.L., 2004. Protecting Personal Information Online : a Survey of User Privacy Concerns and Control Techniques. *Journal of Computer Information Systems*, pp.85–92.

Cho, H., Lee, J.S. & Chung, S., 2010. Optimistic bias about online privacy risks: Testing the moderating effects of perceived controllability and prior experience. *Computers in Human Behavior*, 26(5), pp.987–995. Available at: http://dx.doi.org/10.1016/j.chb.2010.02.012.

Consolvo, S. et al., 2005. Location Disclosure to Social Relations: Why, When, & What People Want to Share. *CHI 2005 Conference on Human Factors in Computing Systems*, pp.81–90.

Culnan, M.J., 1995. Consumer Awareness of Name Removal Procedures: Implications for Direct Marketing. *Journal of Direct Marketing*, 9(2), pp.10–19. Available at: http://doi.wiley.com/10.1002/dir.4000090204.

Culnan, M.J. & Armstrong, P.K., 1999. Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organization Science*, 10(1), pp.104–115.

Das, S. et al., 2014. The Effect of Social Influence on Security Sensitivity. *… Usable Privacy and Security …*, pp.143–157. Available at: http://cmuchimps.org/uploads/publication/paper/147/the_effect_of_social_influence_on_security_sensitivity.pdf.

Dinev, T. & Hu, Q., 2007. The Centrality of Awareness in the Formation of User Behavioral Intention Toward Preventive Technologies in the Context of Voluntary Use. *Journal of the Association for Information Sysytems*, 8(7), pp.386–408.

Dommeyer, C.J. & Gross, B.L., 2003. What consumers know and what they do: An investigation of consumer knowledge, awareness, and use of privacy proteciton strategies. *Journal of Interactive Marketing*, 17(2), pp.34–51. Available at: http://dx.doi.org/10.1002/dir.10053.

Fishbein, M. and Ajzen, I., 1977. Belief, attitude, intention, and behavior: An introduction to theory and research

Fleming, J. & Yu, D. 2015. Consumers doubt their personal info is very safe. [Blog post]. Retrieved from http://www.gallup.com/businessjournal/181904/consumers-doubt-personal-info-safe.aspx

Gefen, D., Karahanna, E. & Straub, D., 2003. Trust and TAM in Online Shopping TRUST AND TAM IN ONLINE SHOPPING: AN INTEGRATED MODEL. *MiS Quarterly*, 27(1), pp.51–90.

Gross, R. & Acquisti, A., 2005. Information revelation and privacy in online social networks. *Privacy in the Electronic Society 2005*, p.11. Available at: http://dl.acm.org/citation.cfm?id=1102214.

Gurung, A., Luo, X. & Raja, M.., 2014. An Empirical Investigation on Customer ' s Privacy Perceptions , Trust and Security Awareness in E-commerce Environment. *Journal of Information Privacy and Security*, 4(1), pp.42–64.

Hill, K., 2012. How Target figured out a teen girl was pregnant before her father did. *Forbes*, 16, pp.2–7.

Hoffman, D.L., Novak, T.P. & Peralta, M., 1999. Building consumer trust online. *Communications of the ACM*, 42(4), pp.80–85.

Horne, D. R., & Horne, D. A. 1998. Domains of Privacy: Toward an Understanding of Underlying Factors. In *Direct Marketing Educators' Conference, San Francisco, CA*

Iachello, G. & Hong, J., 2007. End-User Privacy in Human-Computer Interaction. *Foundations and Trends® in Human-Computer Interaction*, 1(1), pp.1–137. Available at: http://dl.acm.org/citation.cfm?id=1324103.1324104.

ISO 9241-210:2010 Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems

Jai, T.C. & King, N.J., 2015. Privacy versus reward : Do loyalty programs increase consumers ' willingness to share personal information with third-party advertisers and data brokers ? *Journal of Retailing and Consumer Services*, 28(2015), pp.1–8. Available at: http://dx.doi.org/10.1016/j.jretconser.2015.01.005.

Kang, R. et al., 2015. "My Data Just Goes Everywhere:" User Mental Models of the Internet and Implications for Privacy and Security. In *Symposium on Usable Privacy and Security (SOUPS).* pp. 39–52. Available at: https://www.usenix.org/system/files/conference/soups2015/soups15-paper-kang.pdf.

King, J. & Hoofnagle, C.J., 2008. A Supermajority of Californians Supports Limits on Law Enforcement Access to Cell Phone Location Information. *SSRN eLibrary*, (January 2001), pp.1–20. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1137988.

Krane, D., Light, L. & Gravitch, D., 2002. Privacy on and off the Internet: What consumers want. *Harris Interactive*, 10003(15229).

Kumaraguru, P. & Cranor, L., 2005. Privacy indexes: A survey of westin's studies. *School of Computer Science, Carnegie Mellon University*, Tech. rep.(December), pp.1–22. Available at: http://www.casos.cs.cmu.edu/publications/papers/CMU-ISRI-05-138.pdf\nhttp://repository.cmu.edu/isr/856/.

Leino-Kilpi, H. et al., 2001. Privacy:a review of the literature. *International Journal of Nursing Studies*, 38(6), pp.663–671.

Liao, C., Liu, C.-C. & Chen, K., 2011. Examining the impact of privacy, trust and risk perceptions beyond monetary transactions: An integrated model. *Electronic Commerce Research and Applications*, 10(6), pp.702–715. Available at: http://dx.doi.org/10.1016/j.elerap.2011.07.003.

Liu, C. et al., 2005. Beyond concern-a privacy-trust-behavioral intention model of

electronic commerce. *Information and Management*, 42(2), pp.289–304.

Malheiros, M., Preibusch, S. & Sasse, M.A., 2013. "Fairly truthful": The impact of perceived effort, fairness, relevance, and sensitivity on personal data disclosure. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7904 LNCS, pp.250–266.

Malhotra, N.K. et al., 2004. Internet Users' Information Privacy Conernc (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, 15(4), pp.336–355.

Malhotra, N.K., Kim, S.S. & Agarwal, J., 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*, 15(4), pp.336–355.

Margulis, S.T., 1977. Conceptions of Privacy: Current Status and Next Steps. *Journal of Social Issues*, 33(3), pp.5–21. Available at: http://dx.doi.org/10.1111/j.1540-4560.1977.tb01879.x\nhttp://onlinelibrary.wiley.com/doi/10.1111/j.1540-4560.1977.tb01879.x/abstract\nhttp://onlinelibrary.wiley.com.libproxy1.nus.edu.sg/store/10.1111/j.1540-4560.1977.tb01879.x/asset/j.1540-4560.1977.tb018.

Martin, K. & Shilton, K., 2015. Why Experience Matters to Privacy: How Context-Based Experience Moderates Consumer Privacy Expectations for Mobile Applications. *Journal of the Association for Information Science and Technology*.

Masiello, B., 2009. Deconstructing the privacy experience. *IEEE Security and Privacy*, 7(4), pp.68–70.

McKnight, D.H., Choudhury, V. & Kacmar, C., 2002. The impact of initial consumer trust on intentions to transact with a web site: A trust building model. *Journal of Strategic Information Systems*, 11(3–4), pp.297–323.

Morey, T., Forbath, T. & Schoop, A., 2015. CUSTOMER DATA : DESIGNING FOR TRANSPARENCY AND TRUST. *Harward Business Review*, pp.96–105.

Norberg, P.A., Horne, D.R. & Horne, D.A., 2007. The Privacy Paradox : Personal Information Disclosure Intentions vers us Behaviors. *Journal of Consumer Affairs*, 41(I), pp.100–127.

Oulasvirta, A. et al., 2014. Transparency of Intentions Decreases Privacy Concerns in Ubiquitous Surveillance. *Cyberpsychology, Behavior, and Social Networking*, 17(10), pp.633–638. Available at: http://online.liebertpub.com/doi/full/10.1089/cyber.2013.0585.

Patton, M., 1990. Qualitative Evaluation and Research Methods. *Qualitative Evaluation and Research Methods*, pp.169–186. Available at: http://legacy.oise.utoronto.ca/research/field-centres/ross/ctl1014/Patton1990.pdf.

Pavlou, P.A., 2003. Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model. *International Journal of*

Electronic Commerce, 7(3), p.34.

Phelps, J., Nowak, G. & Ferrell, E., 2000. Privacy Concerns and Consumer Willingness to Provide Personal Information. *Journal of Public Policy & Marketing*, 19(1), pp.27–41.

Rose, J., Rehse, O. & Röber, B., 2012. The Vaue of our Digital Identity. *Liberty Global Policy Series, The Boston Consulting Group*. Available at: http://www.lgi.com/PDF/public-policy/The-Value-of-Our-Digital-Identity.pdf.

Schoeman, F. 1984. Privacy: philosophical dimensions. *American Philosophical Quarterly*, *21*(3), 199-213.

Schoenbachler, D.. & Gorden, G.., 2002. Trust and Customer Willingness To Provide Information In Database Driven Relationship Marketing. *Journal of Interactive Marketing*, 16(3), pp.2–16.

SINTEF. 2013, May 22. Big Data, for better or worse: 90% of world's data generated over last two years. *ScienceDaily*. Retrieved October 1, 2016 from www.sciencedaily.com/releases/2013/05/130522085217.htm

Skinner, G., Han, S. & Chang, E., 2006. An information privacy taxonomy for collaborative environments. *Information Management & Computer Security*, 14(4), pp.382–394.

Smith, H.J. et al., 1996. Information Privacy : Measuring Individuals' Concerns about Organizational Practices. *MIS Quarterly*, 20(2), pp.167–196.

Smith, H.J., Dinev, T. & Xu, H., 2011. Information Privacy Research : An Interdisciplinary Review. *MIS Quarterly*, 35(4), pp.989–1015.

Solove, D.J., 2002. Conceptualizing privacy. *California Law Review*, 90(4), pp.1087–1155.

Spiekermann, S., Grossklags, J. & Berendt, B., 2001. E-privacy in 2nd Generation E-Commerce: Privacy Preferences versus actual Behavior. *EC '01 Third ACM Conference on Electronic Commerce*, pp.38–47.

Stone, E.F. et al., 1983. A field experiment comparing information-privacy values, beliefs, and attitudes across several types of organizations. *Journal of Applied Psychology*, 68(3), pp.459–468.

Taylor-Powell, E. & Renner, M., 2003. Analyzing Qualitative Data. *Madison: University of Wisconsin*. Available at: http://www.tandfonline.com/doi/pdf/10.1207/s15430421tip3903_2\nhttps://drive.google.com/file/d/0B2wvlMGjbwrfa3F2MGVNSWk3T1E/edit?usp=sharing.

Warren, S.D. & Brandeis, L.D., 1890. The Right to Privacy. *Harward Law Review*, 4(5), pp.193–220.

Westin, A.F., 1968. Privacy and freedom. *Washington and Lee Law Review*,*25*(1), p.166.

Westin, A.F., 2003. Social and Political Dimensions of Privacy. *Journal of social issues*,

59(2), p.354 S.

Whitley, E. a., 2009. Informational privacy, consent and the “control” of personal data. *Information Security Technical Report*, 14(3), pp.154–159. Available at: http://dx.doi.org/10.1016/j.istr.2009.10.001.

Woodruff, A., Pihur, V. & Consolvo, S., 2014. Would a privacy fundamentalist sell their DNA for $ 1000 ... if nothing bad happened as a result ? The Westin categories , behavioral intentions , and consequences. *… on Usable Privacy and …*, pp.1–18. Available at: https://www.usenix.org/system/files/conference/soups2014/soups14-paper-woodruff.pdf.

Xu, H. et al., 2008. Examining the Formation of Individual’s Privacy Concerns: Toward an Integrative View. *International Conference on Information Systems (ICIS) 2008 Proceedings*, pp.1–16.

Yle uutiset 2016. Retrieved from http://yle.fi/uutiset/3-9055737 November 3rd 2016

Zarsky, T.Z., 2004. Thinking outside the box: considering transparency, anonymity, and pseudonymity as overall solutions to the problems in information privacy in the internet society. *University of Miami Law Review*, 58(1301), pp.1–54. Available at: http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/umialr58&section=57.

Appendix A

Question 1: Who do you trust with your personal information? (information that you can be identified from)

Question 2: Why?

Question 3: What kind of information would you share with the ones you selected above?

Question 4: Age

Question 5: Do you trust your own skills and competence when it comes to keeping personal information safe?

Appendix B

Product attribute questions:

- What size are you looking for?
- Any wishes for color?

Usage oriented questions:

- Where will you use the jacket?
- For what purpose are you getting the jacket?

Personal questions supporting product selection:

- How important is the material of the jacket?
- Do you follow trends?
- Price is the most important feature in the jacket?
- Height?
- Gender?
- Age?

Personal questions independent of product selection:

- How often do you buy a jacket?
- What do you do on your free time?

Appendix C

How would you feel about this type of service?
How did you feel about giving your personal information?

Could you see that this type of application could help you getting better service?
Why did you leave some answers unanswered? Why did you stop there?
Would you be willing to give personal information about yourself in exchange for better personal service?

To whom would you be willing to give your personal information?

How did you experience privacy here? What factors affected that?

Appendix D

1. I would be ready to use the personalized shopping application
2. I felt that the application was safe
3. I feel that application similar to the one I just tested, do not violate my privacy
4. I am willing to give my personal information to companies if I benefit from it
5. I am generally worried about my privacy